Remarks and Dedication: George E.P. Box 1919-2013

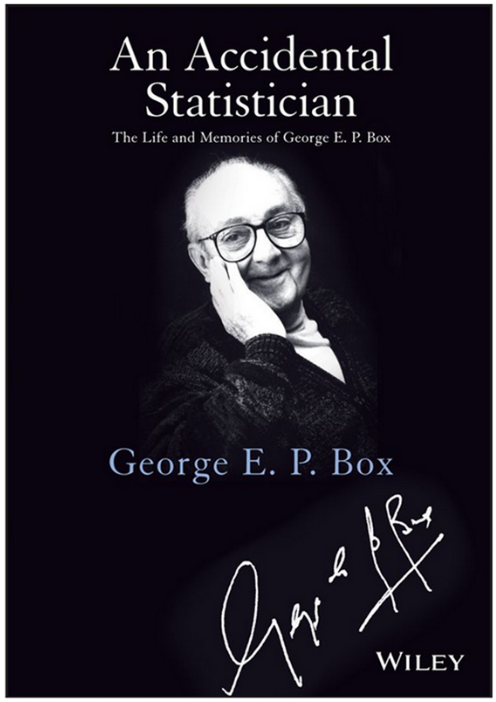
Kevin Little, Ph.D.

15 October 2013

SWQN Annual Meeting

The outline

1. Introduction to George Box
2. Context Questions
3. Welcome to Process World
   – Introduction to Difference Equations
   – Process Models
   – EWMA is “best forecast” …
   – Adjusting to get better performance
4. Context Questions, Answered
5. Appendix 1 Notes and Songs
6. Appendix 2 More on Chapters 4 and 5


An Accidental
Statistician
The Life and Memories of George E. P. Box
George E. P. Box
WILEY

What is Statistics?

"A serious mistake has been made in classifying statistics as part of the mathematical sciences. Rather it should be regarded as a catalyst to scientific method itself."

p. 18 *An Accidental Statistician*

10/14/2013 Informing Ecological Design, LLC• Madison, WI 4

Part of George's early technical education included preparation to acquire an external degree at London University. In the 1930's, applicants needed to pass an Intermediate Science Exam.

"My subjects were pure and applied mathematics, physics (heat, light and sound, electricity and magnetism), and chemistry (organic and inorganic). These were the most difficult exams I ever took, but I passed and they helped me get a grounding in science that has been invaluable ever since."

"I believe that it was this basic scientific knowledge that helped me later on to come up with ideas in the development of statistics. It would, I think, be tremendously helpful if, before taking a degree in statistics, there was a requirement to pass a similar preliminary exam in science. A serious mistake has been made in classifying statistics as part of the mathematical sciences. Rather it should be regarded as a catalyst to scientific method itself. Proper preparation for a degree in statistics should be like that for the intermediate science exam described above, which would include running real experiments." (p. 18)

Based on his work in the Statistical Methods Panel at Imperial Chemical Industries (ICI) with chemists and chemical engineers, George developed Response Surface Methods and Evolutionary Operations.

A few events in George's Life

- 1919 October 18, born in Gravesend England
- 1939 First article published (on experimentation in waste water plants)
- 1939-1945 British Army service, chemical warfare assignment: practical experimentation in England.
- 1946-1949 first degrees in statistics, University College, London
- 1947-1955 work at Imperial Chemical Industries, UK
- 1956-1959 head of Statistical Techniques Research Group at Princeton University
- 1959-1991 Professor of Statistics Department, University of Wisconsin—Madison (founder)
- 1991-2013 Emeritus Professor, world traveler, active author
- 2013 March 23, died in Madison, Wisconsin USA

A few events in George's Life

- 1919 October 18, born in Gravesend England
- 1939 First article published (on experimentation in waste water plants)
- 1939-1945 British Army service, chemical warfare assignment: practical experimentation in England.
- 1946-1949 first degrees in statistics, University College, London
- 1947-1955 work at Imperial Chemical Industries, UK
- 1956-1959 head of Statistical Techniques Research Group at Princeton University
- 1959-1991 Professor of Statistics Department, University of Wisconsin—Madison (founder)
- 1991-2013 Emeritus Professor, world traveler, active author
- 2013 March 23, died in Madison, Wisconsin USA

Book Titles

Statistical Methods in Research and Production
The Design and Analysis of industrial experiments
Evolutionary Operation: A Statistical Method for Process Improvement
Time Series Analysis: Forecasting and Control (four editions)
Bayesian Inference in Statistical Analysis
Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building
(2nd edition: *Statistics for Experimenters: Design, Innovation, and Discovery*)
Empirical Model-Building and Response Surfaces
(2nd edition: Response Surfaces, Mixtures, and Ridge Analyses)
Statistical Control: By Monitoring and Feedback Adjustment (two editions)
An Accidental Statistician: The Life and Memories of George E. P. Box

Collections

The Collected Works of George E.P. Box Volumes 1 and 2 (64 of the major papers through 1982)
Box on Quality and Discovery: With Design, Control, and Robustness
Improving Almost Anything: Ideas and Essays

Book Titles

Work done in 1950s and 1960s

Statistical Methods in Research and Production
The Design and Analysis of industrial experiments
Evolutionary Operation: A Statistical Method for Process Improvement
Time Series Analysis: Forecasting and Control (four editions; first edition published 1970)

Bayesian Inference in Statistical Analysis
Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building
(2nd edition: *Statistics for Experimenters: Design, Innovation, and Discovery*)
Empirical Model-Building and Response Surfaces
(2nd edition: *Response Surfaces, Mixtures, and Ridge Analyses*)
Statistical Control: By Monitoring and Feedback Adjustment (two editions)
An Accidental Statistician: The Life and Memories of George E. P. Box

Collections

The Collected Works of George E.P. Box Volumes 1 and 2 (64 of the major papers through 1982)
Box on Quality and Discovery: With Design, Control, and Robustness
Improving Almost Anything: Ideas and Essays

http://www.wordle.net/ using book title list, weighted by number of editions

But beware: Control is not identical with the basic use of the term in Statistical Process Control! It is related to the engineering control useful in the process world.

MAIN MESSAGE AT SWQN 2013

What we all know is true

- Plot process data in time order, “you can see a lot just by looking”
- Apply control chart rules to distinguish signals of special (assignable) causes from patterns of common cause variation.
- Reduce variation by removing special causes and understanding common causes.
- Don’t adjust a process that is in control, else you will increase variability (*Out of the Crisis*, pp. 327-330, “The funnel experiment”).

But what do we do if…

- Our output process, left to its own devices, will drift away from target?
- Input factors that drive this drift are themselves drifting, not in states of statistical control?
- We can’t technically or economically remove the drifts?

In such a world…

Can we make useful predictions about future performance?

Can we make compensate for the drift in a smart way?

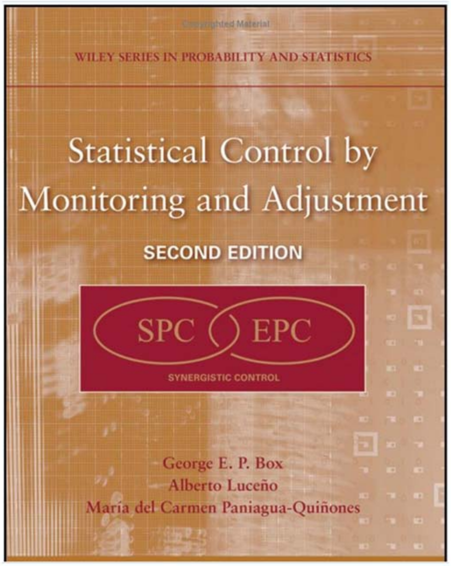
Process World is characterized by flows, flow rates, feedback loops, delays, structural relationships, and (sometimes) factors you can control or adjust to affect the outcomes you want.   This section of the talk is based on chapters 4, 5 and 6 of Statistical Control by Monitoring and Feedback Adjustment, first edition.

Waste water treatment and chemical production of polymers are examples. Biological systems are all in process world.

We'll work with discrete time steps; in many applications, people work with continuous time, too.

WELCOME TO PROCESS WORLD

A guidebook to process world for SWQN



Contents

v

This section of the talk is based on chapters 4-6, first edition (1992).

An overview of chunks of *Statistical Control For Monitoring and Adjustment*

1. The “**Process World”** has features that we can describe with difference equations (Chapter 4)

2. A starting place for process models is a process that wanders “quite a bit” without intervention and adjustment, called the **IMA(1,1)** process (Chapters 1 and 5) (a process “in statistical control” is a special case of IMA(1,1)).

3. An **“Exponentially Weighted Moving Average”** (EWMA) will give you a “best forecast” one step into the future if you are working with the wandering IMA(1,1) process. (Chapter 5)

4. If you have a **knob** that will adjust your IMA(1,1) wandering process, use EWMA to transform the IMA process into a **state of statistical control**. (Chapter 6-9, including issues of cost of adjustment)

Warning!

The next 15 slides have some symbols and equations.

Don't panic, there is no math test today.

Keep breathing and follow the main threads.

INTRODUCTION TO DIFFERENCE EQUATIONS

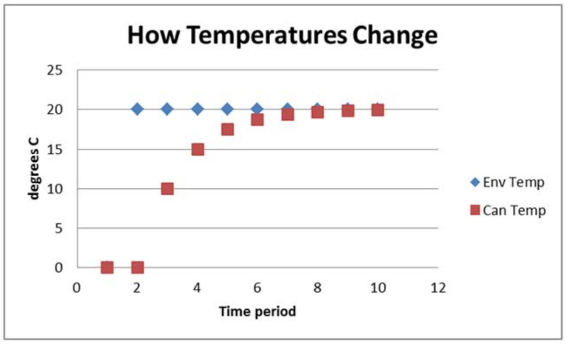
Heat flow: The can of beans

- Can of beans in refrigerator at 0°C.
- Room temperature at 20°C.
- Can will warm to environment temperature at a rate *proportional* to the difference in temperatures at end of time period t.
- Can taken out of refrigerator at end of time period 2.
- At the end of time period 3, can at 10°C.

X_t is the temperature of can's environment
Y_t is the temperature of the can
$1 - \delta$ is the warming rate constant (in the interval (0,1))

p. 89 BL 1st edition

A model for the problem



The conditions on the previous slide lead to:
$Y_{t+1} - Y_t = (1\text{-}\delta)(X_t - Y_t)$
and $1\text{-}\delta = 0.5$

t	X_t	Y_t
1	0	0.0
2	20	0.0
3	20	10.0
4	20	15.0
5	20	17.5
6	20	18.8
7	20	19.4
8	20	19.7
9	20	19.8
10	20	19.9

Manipulating the Difference Equation

$Y_{t+1} - Y_t = (1-\delta)(X_t - Y_t)$ ⇨ $Y_{t+1} = (1-\delta)X_t + \delta\, Y_t$

But $Y_t = (1-\delta)X_{t-1} + \delta\, Y_{t-1}$

and $Y_{t-1} = (1-\delta)X_{t-2} + \delta\, Y_{t-2}$

and so on…

Substitute the Y expressions, back in time, to get:

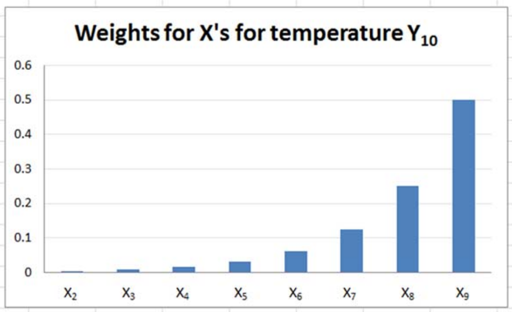
$$Y_{t+1} = (1-\delta)(X_t + \delta X_{t-1} + \delta^2 X_{t-2} + \delta^3 X_{t-3} + \ldots + \delta^{t-1} X_o) + \delta^t Y_o$$

⇨ $Y_{t+1} = \widetilde{X}_t$

Info about X's at time t determines Y_{t+1}

(for t large, we can ignore $\delta^t Y_o$)

$\widetilde{X}_t$ is an Exponentially Weighted Moving Average (EWMA)



1. $\widetilde{X}_t$ is a sum of X's
2. A simple average of n items sums each item with weight 1/n.
3. Our sum has exponential weights like (1-δ) δ^0, (1-δ) δ^1, (1-δ) δ^2 etc.
4. As we step in time, the weights move, too
5. It forecasts one step ahead, $\hat{Y}_{t+1} = \widetilde{X}_t$

For *any set of numbers* $\{z_t\}$ and *any value* θ in (-1,1)

Let $\hat{z}_{t+1} = \tilde{z}_t = \lambda(z_t + \theta\, z_{t-1} + \theta^2\, z_{t-2} + \ldots)$ with $\lambda = 1 - \theta$

i.e. the <u>forecast</u> of z at time t+1 is an EWMA(θ)

and define $e_t = z_t - \hat{z}_t$

i.e. define e_t as the <u>forecast error</u> at time t

THEN (among other useful relationships*)

$$z_t - z_{t-1} = e_t - \theta\, e_{t-1}$$

*see Table 4.2 BL 1st edition p. 97 reproduced in Appendix 2

Interim summary

Difference equations capture many useful features of process world.

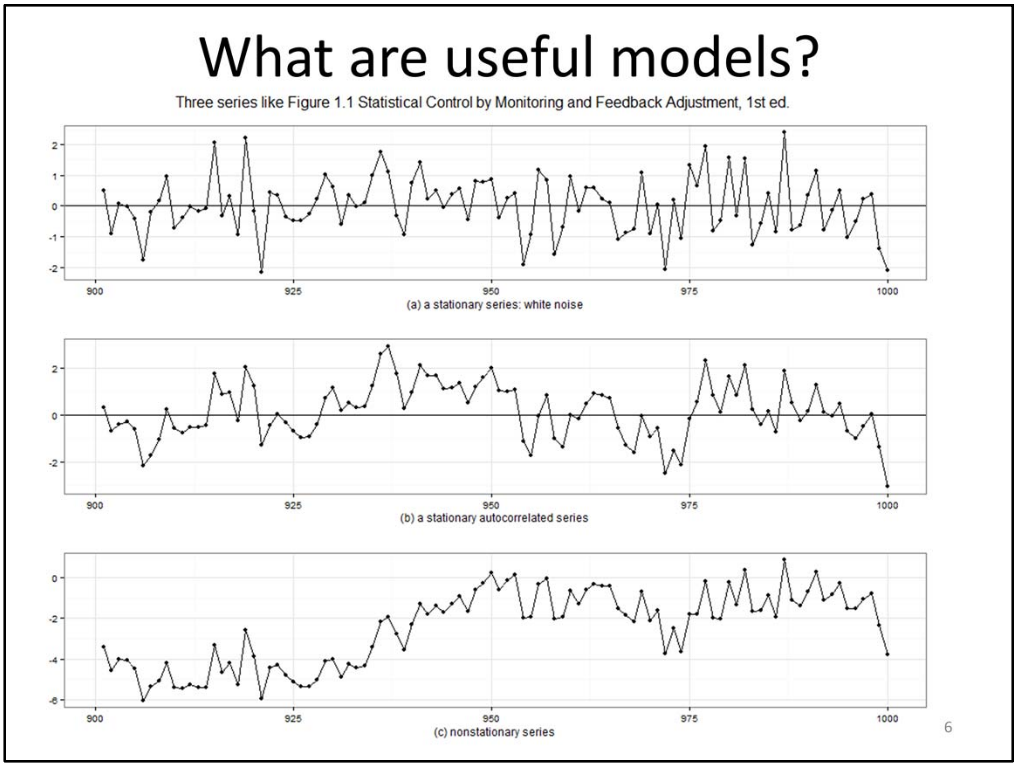
If a difference equation model leads to an EWMA, you have a tool to predict the output, one step ahead.

So far, we have not introduced any statistical variation (common or special causes).

PROCESS MODELS: VARIATION

10/14/2013 Informing Ecological Design, LLC• Madison, WI 25

Stationary models: we have mathematically defined mean and variance for the series (related to convergent and divergent series)




```
#Figure 1.1 example 12 October 2013 code in R
require(ggplot2)
set.seed(1245)
y<-rnorm(1000,0,1)
y2<-vector(mode="numeric",length=1000)
y2[1]<-0
alpha<-0.7
for(i in 2:1000){
  y2[i]<-alpha*y2[i-1]+y[i]
}
summary(y2)

seq<-c(1:1000)
plot(seq,y)
plot(seq,y2)
ynew<-y[901:1000]
ynew1<-y2[901:1000]
seq1<-seq[901:1000]

df1<-as.data.frame(cbind(seq1,ynew))
df2<-as.data.frame(cbind(seq1,ynew1))
p1<-
ggplot(df1,aes(seq1,ynew))+geom_line()+theme_bw()+geom_hline(aes(yintercept=0))+geo
```

```
m_point()+ylab("")+xlab("(a) a stationary series")
p2<-
ggplot(df2,aes(seq1,ynew1))+geom_line()+theme_bw()+geom_hline(aes(yintercept=0))+geo
m_point()+ylab("")+xlab("(b) a stationary autocorrelated series")

y0<-0
z<-vector(mode="numeric",length=1000)
theta<-0.5
z[1]<-0
for(i in 2:1000){
  z[i]<-y[i]-theta*y[i-1]
}
yIMA<-vector(mode="numeric",length=1000)
yIMA[1]<-0
for(i in 2:1000){
  yIMA[i]<-yIMA[i-1]+z[i]
}
summary(yIMA)
yIMA2<-yIMA[901:1000]
summary(z)
df3<-as.data.frame(cbind(seq1,yIMA2))
p3<-ggplot(df3,aes(seq1,yIMA2))+geom_line()+theme_bw()+geom_point()+ylab("")+xlab("(c)
nonstationary series")
grid.arrange(p1,p2,p3, main=("Three series like Figure 1.1 Statistical Control by Monitoring
and Feedback Adjustment, 1st edition"))
```

Non-stationary series (c): An Integrated Moving Average

(c) is an example of an integrated moving average, called an "IMA(1,1)" with $\theta = 0.5$

The difference equation form for IMA(1,1):

$$z_{t+1} - z_t = a_{t+1} - \theta a_t$$

where $\{a_t\}$ is a white noise series.

Our terms really are consistent with what we have already said but would take us a few minutes to get there….Here's a start:

The moving average part of the name comes from the same kind of substitutions we used on slide 17 now applied to $w_{t+1} = a_{t+1} - \theta a_t$ to get a moving average of w values. The integration comes because the "un do" operation of differencing goes by the name of integration (works same as in calculus) and we will work with not the differenced series $\{z_{t+1} - z_t\}$ but the integrated series $\{z_t\}$.

The form has a single difference in the z series and single difference in the error series. Hence the (1,1) after the IMA abbreviation.

Whew.

More jargon: white noise is typically represented by a sequence of a realization of independent and identically distributed normal random variables; mean 0 and variance sigma-squared.

Two arguments for non-stationary models

1. Theory--2nd Law of Thermodynamics:
"left to itself, the entropy (or disorganization) of any system can never decrease." p. 1 1st edition BL

2. Practice--Control methods have been used by engineers for centuries.
Control methods transform non-stationary disturbances to stationary series (like (a) white noise or (b) auto-correlated series.)
p. 5 1st edition BL

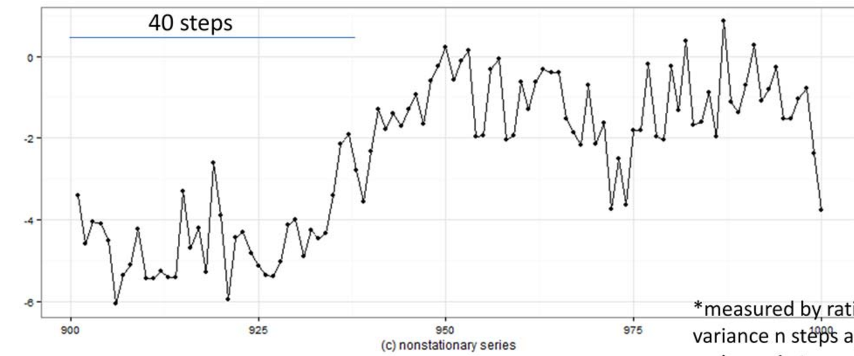
10/14/2013 Informing Ecological Design, LLC• Madison, WI 28

"If left to themselves, machines to not stay adjusted, components wear out, and managers and operators forget, miscommunicate, and change jobs. Thus a stable stationary state is an unnatural one and its achievement requires a hard and continuous fight. It is hard because it requires us to try and undo the effects of the 2nd law of thermodynamics. This law says that, *left to itself*, the entropy (or disorganization) of any system can never decrease." p. 1 1st edition BL

"…to design an efficient feedback control scheme we need to know something about the disturbance we are trying to control—that is, the sequence of deviations that would have occurred if not control had been applied." Such a disturbance can frequently be represented by a nonstationary series, such as that shown in Figure 1.1c." p. 5 BL 1st edition

Basic Property of IMA(1,1)

the variability* between observations grows linearly. In the example, with $\theta = 0.5$, the standard deviation 40 steps apart is 3 times the standard deviation one step apart:



*measured by ratio of variance n steps apart to variance 1 step apart

10/14/2013 Informing Ecological Design, LLC• Madison, WI 29

And also see discussion in section 5.5, interpretation of IMA(1,1) as sticky innovation generating model due to Box and Kramer (1992).

EWMA GIVES BEST FORECAST FOR IMA(1,1) MODEL

Four Step Logic

1. If the one-step forecast errors make a "white noise" series, there's no way to improve forecasts internal to our forecasting system. We've got the "best" forecast.
2. For *any set of numbers* $\{z_t\}$ where we set up an EWMA, we get $z_t - z_{t-1} = e_t - \theta\, e_{t-1}$ where e_t is the forecast error $z_t - \hat{z}_t$
3. IMA(1,1) has form: $z_{t+1} - z_t = a_{t+1} - \theta a_t$, $\{a_t\}$ white noise
4. Given IMA, set up EWMA. Forecast error is now a white noise series and you have the "best forecast."

It turns out even if the disturbance series is NOT an IMA(1,1) the EWMA set up still gives "surprisingly good forecasts." The math details are in E.J. Muth (1960), "Optimal properties of exponentially weighted forecasts of time series with permanent and transitory components," *Journal of the American Statistical Association*, **55**, 299-306.

ADJUSTING TO GET BETTER PERFORMANCE

Adjusting a process--Set up

- Adjustment knob position at time t is X_t
- One unit of adjustment produces **g** units change in the output.
- Deviations from target (disturbances) are $\{z_t\}$

If we could look into the future, we would adjust the process to exactly compensate the disturbance:

$$gX_t = - z_{t+1}$$

A very good alternative is:

$$gX_t = \hat{z}_{t+1} = \tilde{z}_t \quad \text{(our friend, EWMA)}$$

10/14/2013 Informing Ecological Design, LLC• Madison, WI 33

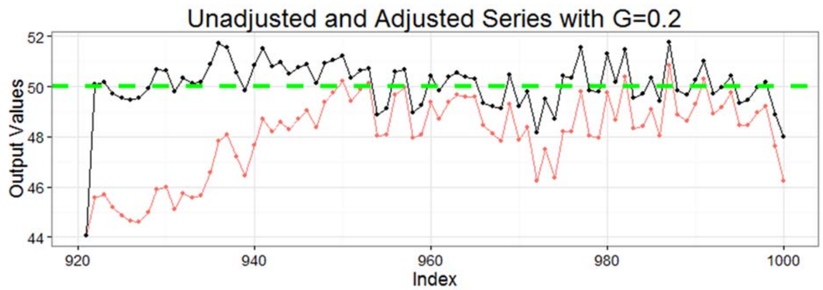
The adjustment scheme is **optimal** for an IMA(1,1) series because EWMA provides “best” forecast as explained qualitatively by BL in chapter 5.

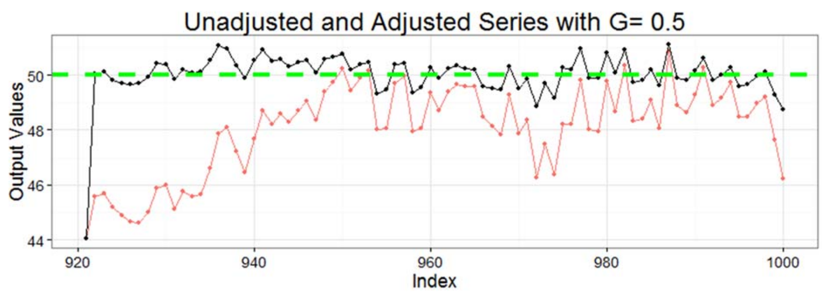
Adjusting a process--Example

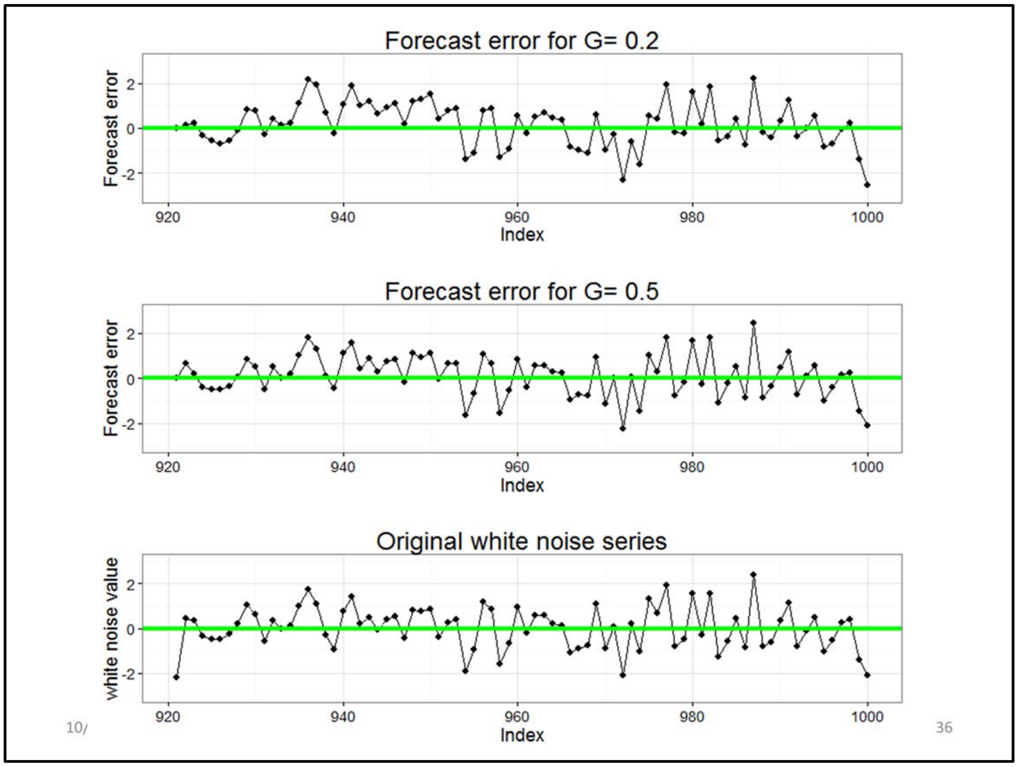
1. The disturbance process is IMA(1,1) with θ = 0.5
2. We have an adjustment knob, with g = 1.
3. The target is **50**
4. We use the EWMA technology to make adjustments
   a. with factor 0.2 (which would be "best" for θ = 0.8)
   b. with factor 0.5, the perfect adjustment, because it matches θ = 0.5.

If you didn't know anything for sure about the EWMA factors, you could try 0.2. It's pretty good, for many reasons discussed in Chapter 6.

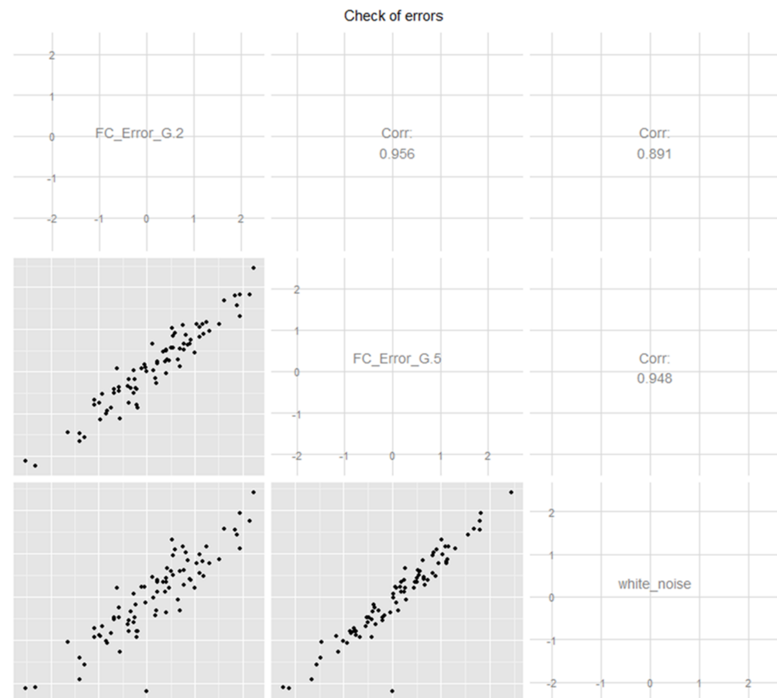
You can to all of this work with spreadsheets, which are great for adding up and differencing. We continued to work in R.


Unadjusted and Adjusted Series with G=0.2
Output Values
Index
44
46
48
50
52
920
940
960
980
1000


Unadjusted and Adjusted Series with G= 0.5
Output Values
Index
44
46
48
50
920
940
960
980
1000



Forecast error for G= 0.2
Forecast error
Index
Forecast error for G= 0.5
Forecast error
Index
Original white noise series
white noise value
Index
10/
36


Check of errors
FC_Error_G.2
Corr:
0.956
Corr:
0.891
FC_Error_G.5
Corr:
0.948
white_noise

SUMMARY

In much of the real world…

- Our output process, if left to its own devices, will drift away from target.
- Input factors that drive this drift are themselves drifting, not in states of statistical control.
- We often can't technically or economically remove the drifts.

In such a world…

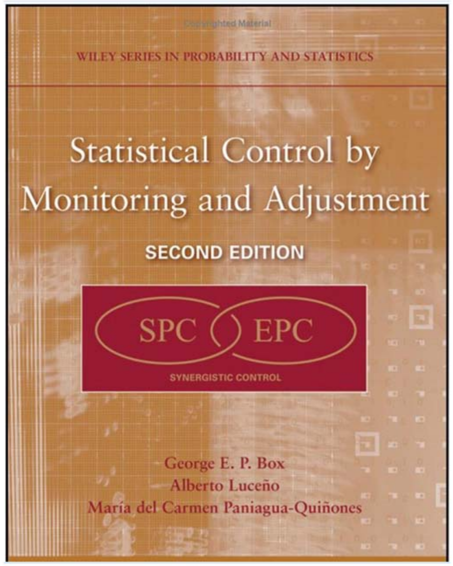
Can we make useful predictions about future performance?

Yes

Can we compensate for the drift in a smart way?

Yes, if we have an adjustment knob

I recommend this guidebook to process world!



APPENDIX 1: NOTES AND SONGS

New book* has the reference right

*_Transforming Health Care Leadership: A Systems Guide to Improve Patient Care, Decrease Costs, and Improve Population Health_ by M. Maccoby, C. L. Norman, C.J. Norman, R. Margolies (2013), Jossey-Bass, San Francisco, location 1625 Kindle edition)

Context: using polynomial functions to approximate an unknown smooth function in a region of interest $R(\boldsymbol{\xi})$, where $\boldsymbol{\xi}$ is a set of variables, with some function $\eta(\boldsymbol{\xi})$ that relates the response y to the variables.

"...writing $f(\boldsymbol{\xi})$ for the polynomial approximation, we have

$$E(y) = \eta(\boldsymbol{\xi})$$

for all $\boldsymbol{\xi}$, and

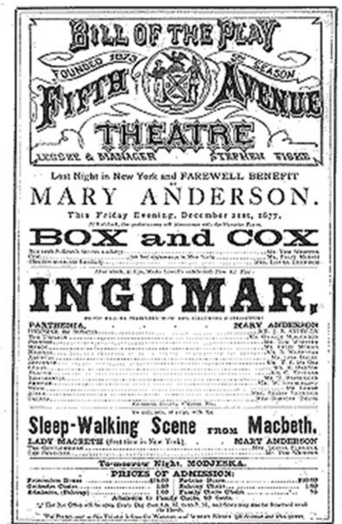
$$\eta(\boldsymbol{\xi}) \approx f(\boldsymbol{\xi})$$

over some limited region of interest $R(\boldsymbol{\xi})$. The fact that the polynomial is an approximation does not necessarily detract from its usefulness because all models are approximations. Essentially, **all models are wrong, but some are useful.** However the approximate nature of the model must always be borne in mind."

Empirical Model Building and Response Surfaces (1987) Wiley, p. 424

Why did Box and Cox write a paper?

George Box and David Cox (later, Sir David) both served on the Research Committee of the Royal Statistical Society in the 1950's. Their names caused their colleagues to tease them, referring to a 1847 play called *Box and Cox* (revised as Cox and Box in the late 1860s, the first successful comic opera by Arthur Sullivan.)

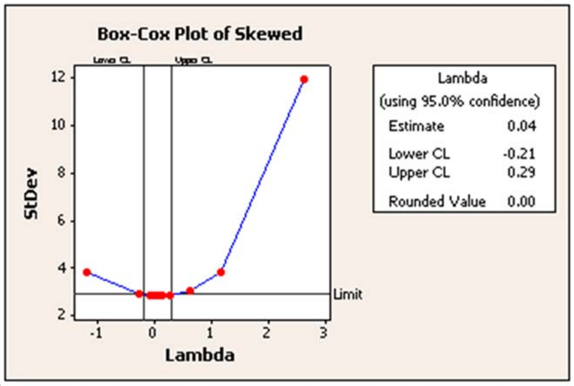


http://en.wikipedia.org/wiki/File:CoxBoxNY.gif

The story is about a pair of boarders who rent the same room without initial knowledge of the other (the landlady makes twice the income), as one works the day shift and one works the night shift.
"Box and Cox" is defined by the Oxford English Dictionary as "used to refer to an arrangement whereby people make use of the same accommodation or facilities at different times, according to a strict arrangement"

To settle the teasing, Box and Cox decided to write a paper about transformations, to keep with the story. "In this spirit, we deliberately included two different derivations of our results, one using likelihood and the other Bayes. At the [RSS Research] meeting, the discussants tried hard to find our who wrote what, but were not about to tell them. From a practical point of view, it didn't make much difference anyway." (p. 56)

The Box and Cox legacy



The original paper (JRSS, series B, 26, 2 (1964),211-252) will require you to understand and appreciate a number of technical details. For most of us, the lasting implications of the paper are given by software implementations (e.g. Minitab 16). The transformation that minimizes the test sum of squares will tend to yield a set of transformed observations with approximate normality, for which many of the classic techniques are effective and appropriate.

Statistics and the Environment, 40 years ago, guides us today

"Opinions differ as to how long it will take before various predictable [environmental] crises occur and how much each problem will complicate the solution of the others, but it is very clear that we will be hard pressed and we will be lucky to escape by the skin of our teeth. The truth is that although we are called on to meet very difficult problems of great urgency we know pathetically little of the facts. So we must learn fast."

G.E.P. Box, "Statistics and the Environment", *Journal of the Washington Academy of Sciences*, No. 2, 1974, p. 52

Beer and Statistics are partially confounded. Why?-1

Every student of statistics learns about the t-distribution. W.S. Gossett (1876-1937) discovered that distribution as part of his 38-year career at Guinness, at which his last position was Head Brewer. Gossett published his statistical papers using a pseudonym "Student" to conform with company policy on intellectual property—though what Student invented is not what students of introductory statistics typically learn.

10/14/2013 Informing Ecological Design, LLC• Madison, WI 47

Gossett, trained in chemistry and mathematics, asked and answered many useful questions about the economics of brewing, guided by data analysis and experimentation. He sought help from the principal statisticians of his time, including Karl Pearson and R.A. Fisher. Nevertheless, he knew enough to keep his wits about him and keep his focus on problems of prediction and brewing economics without being ensnared in Pearson-Fisher philosophical controversies. Stephen Ziliak provides an insightful summary of important aspects of Gossett's work in "Guinessometrics: The Economic Foundation of 'Student's' t" (2008), published in the Journal of Economic Perspectives and available at http://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.22.4.199

Beer and Statistics. Why?-2

George sponsored a Beer and Statistics seminar for 30 years at the University of Wisconsin, starting in 1961. Monday nights during the academic year, George invited scientists or engineers to present interesting problems that faculty and students could then discuss, led by George. Refreshments included beer, to help make the sessions less formal than the typical university seminar.

1980 marked the low point of the number of brewers in the U.S. (aside from prohibition years). Point from Stevens Point, WI and Leinenkugel from Chippewa Falls were cheap and local to Madison

10/14/2013 Informing Ecological Design, LLC• Madison, WI 48

George was a masterful consultant--he liked to help scientists and engineers make progress using statistical methods and his command of relevant statistical theory and methods moved the conversation along briskly. I learned a lot about statistical consulting by attending those sessions.

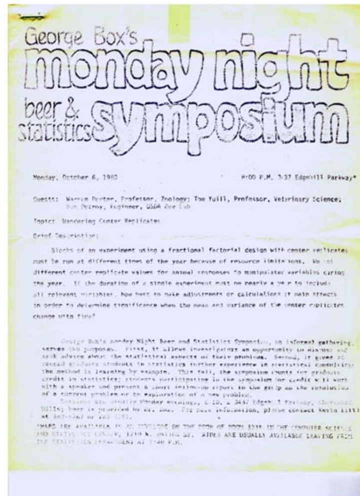
In 1980 and 1981, I had the job to find guest presenters for the Beer and Statistics seminar. Gossett would have been a perfect guest--he had scientific and commercial problems that were challenging and he appreciated the power of proper analysis and experimentation that George could describe.

Gossett's statistical approach to brewing, coupled with more modern methods of inference and data display could benefit the new generation of brewers in North America.

The Beer Seminar

Here's part of the advertising copy for the seminar from 1980:

George Box's Monday Night Beer and Statistics Symposium, an informal gathering, serves two purposes. First, it allows investigators an opportunity to discuss and seek advice about the statistical aspects of their problems. Second, it gives advanced graduate students in statistics further experience in statistical consulting; the method is learning by example. (Thanks to Dr. Conrad Fung for digging this information out and sharing with me.)



George Box's
monday night
beer & statistics symposium

Monday, October 6, 1980 8:00 P.M. 3-37 Edgehill Parkway*

Guests: Warren Porter, Professor, Zoology; Tom Yuill, Professor, Veterinary Science; [illegible], Engineer, USDA [illegible] Lab

Topic: Wandering Center Replicates

Brief Description:

Blocks of an experiment using a fractional factorial design with center replicates must be run at different times of the year because of resource limitations. We get different center replicate values for animal responses to manipulated variables during the year. If the duration of a single experiment must be nearly a year to include all relevant variables, how best to make adjustments or calculations of main effects in order to determine significance when the mean and variance of the center replicates change with time?

George Box's Monday Night Beer and Statistics Symposium, an informal gathering, serves two purposes. First, it allows investigators an opportunity to discuss and seek advice about the statistical aspects of their problems. Second, it gives advanced graduate students in statistics further experience in statistical consulting; the method is learning by example. This fall, the symposium counts for graduate credit in statistics; students participating in the symposium for credit will work with a speaker and present a short follow-up report to the group on the resolution of a current problem or an exploration of a new problem.

[illegible] Monday evenings, 8:00, [illegible] Edgehill Parkway, [illegible] Hills; beer is provided by Dr. Box. For more information, please contact Kevin Little at [illegible].

[illegible] ARE AVAILABLE [illegible] ROOM [illegible] IN THE COMPUTER SCIENCE AND STATISTICS CENTER, 1210 W. DAYTON ST. RIDES ARE USUALLY AVAILABLE LEAVING FROM THE STATISTICS DEPARTMENT AT [illegible] P.M.

"I don't know why you say goodbye, I say hello"

George was an amateur musician who had a good musical ear and a remarkable memory for lyrics of all kinds.

One day ~32 years ago, George told me that he had become interested in Lennon and McCartney tunes because of their clever use of the circle of fifths*. He looked for and found interesting patterns everywhere!

*"…to music as the periodic table of elements is to chemistry."
http://www.ultimate-guitar.com/lessons/the_basics/the_circle_of_fifths_music_theory_for_dummies.html?no_takeover

Cogent Advice from George: It ain't necessarily so….

Recorded 28 October 1983 Madison, WI in the basement of Professor of Engineering Mac Berthouex, Madison, WI.

Adapted from George and Ira Gershwin's song of the same name in *Porgy and Bess* (1935)

Now life testing theories can show
Those life testing theories can show
That if it's reliable, the distribution's a Weibull
But it ain't necessarily so.

Now Bayes solves all problems, we know.
Yes Bayes solves all problems we know.
We follow that Gospel, so far as it's poss'ble
But it ain't necessarily so.

Now science knows no status quo
But it iterates to and fro
Inductive, deductive, the theory's seductive
But it ain't necessarily so.

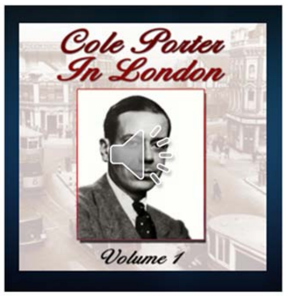
ARIMA we learned long ago
ARIMA we learned long ago
Can answer all queries regarding time series
But it ain't necessarily so.

Analysis should be robust
In these days robust is a must
Outliers should be fated to all be downweighted
But it ain't necessarily so.

Final reminder from George, to do what all good scientists do…

Cole Porter (1933) "Experiment" from *Nymph Errant;* lyrics reprinted with permission in both editions of *Statistics for Experimenters* (Wiley, 1978 and 2005)

Performed at George's memorial celebration 5 May 2013, Madison, WI and the closing words in his printed obituary



You can buy this version from Amazon

Experiment!
Make it your motto day and night.

Experiment,
And it will lead you to the light.

The apple on the top of the tree
Is never too high to achieve,
So take an example from Eve…
Experiment!
Be curious,
Though interfering friends may frown.
Be furious
At each attempt to hold you down.

If this advice you only employ
The future can offer you infinite joy
And merriment…
Experiment.
And you'll see!

Obituary Box, George E. P. "Pel"

APRIL 07, 2013 5:00 AM SHOREWOOD HILLS - George E. P. "Pel" Box, age 93, died on Thursday, March 28, 2013, at home. He was much loved and will be missed by many. A celebration of his life will be held at 2 p.m. on Sunday, May 5, 2013, at FIRST UNITARIAN SOCIETY, 900 University Bay Drive, Madison.

Pel was born in Gravesend, Kent, England in 1919. Born into modest means, both he and his brother, Jack earned scholarships in order to attend the more elite public school in Kent. Pel began his scientific life as a chemist, publishing his first paper at the age of 19 on the activated sludge process to produce clean effluent. Understanding the dangers of fascism, he abandoned his education and enlisted in the army on the very first day he was eligible. During his six years in the army he eventually was sent to Porton Down Experimental Station to study the potential impact of poison gases. George realized that only a statistician could get reliable results from experiments so he taught himself statistics and a career was born.

After the war, George worked at Imperial Chemical Industries for eight years, spending 1953, however, at North Carolina State where he met some of the pre-eminent statisticians of the day. After working again in England for a few years, he left ICI and went to Princeton University in 1956 to direct a statistical research group. George came to Madison in 1959 establish the University of Wisconsin's Department of Statistics in 1960 and retired as an emeritus professor in 1992, though he continued to contribute research papers and write books until his death.

George co-founded the UW Center for Quality and Productivity Improvement with William "Bill" Hunter in 1985. He wrote and co-authored major statistical books on evolutionary operation, times-series, Bayesian analysis, the design of experiments, statistical control, and quality improvement. George loved his students and was proud all their contributions. His last
book, a memoir called, An Accidental Statistician: The Life and Memories of G.E.P. Box, is soon to be published by Wiley.

George received many honors, including being elected Vilas Research Professor of Statistics at the UW Madison in 1980, and a Fellow of the Royal Society, in 1985. George is remembered for his long-running Monday Night Beer and Statistics sessions, held at his house and open to all campus researchers and students. He provided analysis and suggestions, and many students later said that they learned more from these sessions than from any classroom or textbook.

Pel had encyclopedic knowledge of song lyrics and almost every poem or verse he had every heard. His favorite book, however, was Alice in Wonderland. We also remember his skits and songs such as "There's No Theorem Like Bayes' Theorem." We treasure him for his wit, modesty, kindness and warmth; his love of family and friends.

He is survived by his wife, Claire Quist Box; his former wife, Joan Gunhild Box; their daughter, Helen Elizabeth Box (Tom Murtha) and children, Isaac and Andy of Oak Park, IL; son, Harry Christopher Box (Stacey) and children, Henry and Eliza of Cummington, Mass.; son, Simon (Wendy) and children, Mark, Emma, and great-granddaughter, Olivia. He is also survived by the loving family of his late brother, Jack; Michael (Angela) and children, Sarah and Timothy; Roger (Jean); Margaret (Kevin) and son, James; and Dana (John) and children, Katie and Charlie.

A special thanks to Agrace HospiceCare in Madison for their steadfast and kind support. Heartfelt thanks for the loving care provided by Kolleen Bakkum and 3 year old Lincoln and deep gratitude to Gary (Frankie) Johnson for his many years of friendship and support.

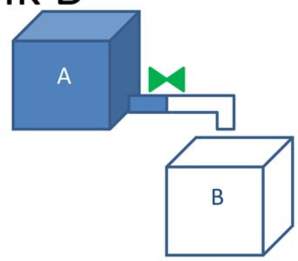
George was buried on April 2, 2013, at the Farley Center for Peace, Justice and Sustainability in the Natural Path Sanctuary. If you want to honor the memory of George, contributions could be made to UW Foundation-George Box Endowment Fund (for the support of graduate students) US Bank Lock Box 78807, Milwaukee, WI 53278; Agrace HospiceCare, 5395 E. Cheryl Parkway Madison, WI 53711.

And a last message from George, "Experiment! Make it your motto day and night. Experiment, And it will lead you to the light …Be Curious,…Get Furious… Experiment, And you'll see!" Cole Porter

APPENDIX 2: CHAPTERS 4 AND 5 BL 1ST EDITION NOTES

Process World: Water flowing from tank A to tank B

- Two identical tanks A and B, cubes with side 100 cm.
- Rate of flow into B proportional to the head (depth) in tank A
- At time 0, tank A is full, tank B is empty.
- At end of time step 1, the valve is opened.
- At the end of time step 2, tank B is filled to 20 cm.
- How do the depths of the tanks vary over time?



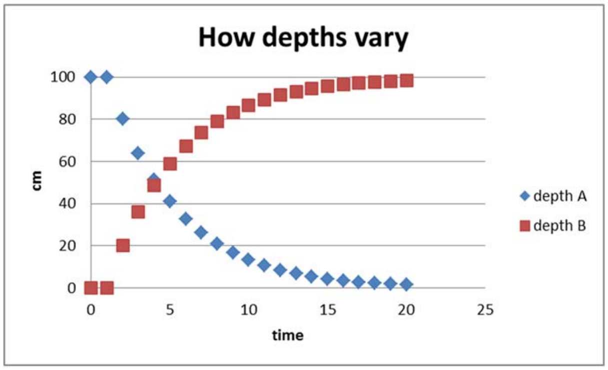
A_t =depth of tank A at end of time period t; B_t = depth of tank B at end of time period t.

Exercise 4.1 BL 1st edition p. 89

A model for the problem (assuming you don't fiddle with the valve!)

The conditions on the previous slide lead to:

$B_{t+1} = 0.2A_t + B_t$



t	A_t	B_t
0	100	0
1	100	0
2	80	20
3	64	36
4	51.2	48.8
5	41.0	59.0
6	32.8	67.2
7	26.2	73.8
8	21.0	79.0
9	16.8	83.2
10	13.4	86.6

A “difference” equation arises…

because we can rewrite our equation

$$B_{t+1} = 0.2A_t + B_t$$

$$\Rightarrow B_{t+1} - B_t = 0.2A_t$$

to describe the **difference** in tank B depth over one time interval

Table 4.2 BL 1st edition p. 97

Important Relationships for Exponentially Weighted Moving Averages (which hold for any series of numbers $\{z_t\}$ and any value of θ in (-1,1)—though the practical values of interest are usually in (0,1)):

Let $\hat{z}_{t+1} = \tilde{z}_t = \lambda(z_t + \theta z_{t-1} + \theta^2 z_{t-2} + \ldots)$ with $\lambda = 1 - \theta$ and $e_t = z_t - \tilde{z}_{t-1} = z_t - \hat{z}_t$

Algebra	Words
A. $\tilde{z}_t = \lambda z_t + \theta\tilde{z}_{t-1}$	The EWMA at time t is an interpolation between the current observation and the previous EWMA
A1. $\hat{z}_{t+1} = \lambda z_t + \theta\hat{z}_t$	The one step ahead forecast at time t is an interpolation between the current observation and the previous forecast
B. $\tilde{z}_t - \tilde{z}_{t-1} = \lambda(z_t - \tilde{z}_{t-1})$	The EWMA difference is a fraction of the difference between the current observation and the previous EWMA.
B1. $\hat{z}_{t+1} - \hat{z}_t = \lambda e_t$	The difference in forecasts is a fraction of the forecast error.
C. $z_t - z_{t-1} = e_t - \theta e_{t-1}$	The difference in observations is a simple function of the forecast errors.

The Variogram

"If left to themselves, machines continue to go out of adjustment, tools continue to wear out, and people continue to miscommunicate and forget. In all such circumstances, it is reasonable to expect that, if no corrective action is taken, observations spaced further apart will differ more and more."

BL 1st edition, p. 114

Let V_m = variance($z_{t+m} - z_t$)

In particular,

V_1 = variance($z_{t+1} - z_t$)

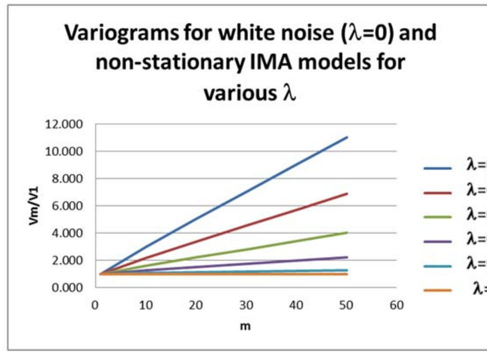
Variogram:

plot of V_m / V_1 versus m

Variogram Insight: IMA(1,1) is the simplest model with increasing V_M/V_1

"…a realistic model [for uncontrolled disturbances] would be one in which a steady increase in V_m/V_1 occurred as m increased."

V_m/V_1 increasing linearly in m *corresponds exactly* to the IMA(1,1) non-stationary model. BL 1st edition, p. 114



See figure 5.3 BL 1st edition, p. 115

Contact information

Kevin Little, Ph.D.
klittle@iecodesign.com
608-251-4355