A Close Look at Run Length Distribution

Dennis Lin
University Distinguished Professor
Department of Statistics
The Pennsylvania State University

At
NCTS
National Tsing Hua University
21 December, 2012

2012.12.21

世界末日

2012.12.22

Vincent Lin’s Birthday

---

14 Dec, 2012

林子璿(中)領獎，旁為男女主角。

NBC頒發最佳導演大獎
華裔林子璿掄元獲6萬

本報記者王鏑紐約報道

NBC電視台10月舉行第六屆短片電影頒獎典禮(Short Cuts Film Festival)，聯問華裔導演林子璿(Vincent Lin)從超過1540位參賽者中脫穎而出，最終掄元，更獲得6萬元獎金。

這條大約15分鐘的短片名為「愛·紐約」，林子璿昨(13日)說，片中利用一位新移民與本地人的故事，訴述在語言無用之下真愛如何帶動溝通，培養出一段真摯感情。這段短片最終受到評委青睞，林子璿也因此獲頒最佳導演獎。目前觀眾登陸youtube網站及輸入http://youtu.be/4GRsW6a7Jjw，便能欣賞。

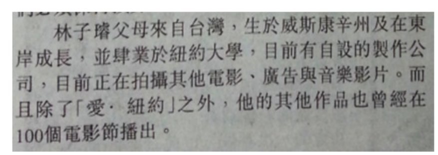
林子璿並介紹說，劇中使用的是英語及韓文，主要是因為試鏡時剛好韓裔的男女主角在現實中也是夫婦，他們對於情感的交流恰到好處，而且男角Randall Park本身已參演過其他電影，有一定知名度。另外，電影場景主要都是在室內，他說那是因為冬天氣溫過冷，而且在室外拍攝成本也較高，他們必須保持預算。

林子璿父母來自台灣，生於威斯康辛州及在東岸成長，並肄業於紐約大學，目前有自設的製作公司，目前正在拍攝其他電影、廣告與音樂影片。而且除了「愛·紐約」之外，他的其他作品也曾經在100個電影節播出。

林子璿(中)領獎，旁為男女主角。

NBC頒發最佳導演大獎
華裔林子璿掄元獲6萬

本報記者王鏑紐約報道



林子璿父母來自台灣，生於威斯康辛州及在東岸成長，並肄業於紐約大學，目前有自設的製作公司，目前正在拍攝其他電影、廣告與音樂影片。而且除了「愛·紐約」之外，他的其他作品也曾經在100個電影節播出。

Dimensional Analysis & Its Applications in Statistics

Dennis Lin
Department of Statistics
The Pennsylvania State University

at
NCTS Workshop on Industrial Statistics
National Tsing Hua University
27 June, 2012

It is unlikely that your paper will be published, if you write

$$1 + 1 = 2$$

It is most likely that your paper will be published, if you write

$$\ln[\lim_{y\to 0}(1+\frac{1}{y})^{y}]+\int_0^\infty \frac{1}{\beta^\alpha\Gamma(\alpha)}y^{\alpha-1}e^{-\frac{y}{\beta}}dy = \sum_0^\infty \frac{\cosh(y)\sqrt{1-\tanh^2(y)}}{2^n}$$

$$1 + 1 = 2$$

$$\ln e$$

$$\int_0^\infty \frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}}dx$$

$$\sum_0^\infty \frac{1}{2^n}$$

$$1 = \cosh(y)\sqrt{1-\tanh^2(y)}$$

$$e = \lim_{z\to 0}(1+\frac{1}{z})^z$$

Now, put them together

$$\ln[\lim_{z\to 0}(1+\frac{1}{z})^{z}]+\int_0^\infty \frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}}dx = \sum_0^\infty \frac{\cosh(y)\sqrt{1-\tanh^2(y)}}{2^n}$$

The Evolution of Science

- **Observational Science**
  - Scientist gathers data by direct observation
  - Scientist analyzes data
- **Analytical Science**
  - Scientist builds analytical model
  - Makes predictions.

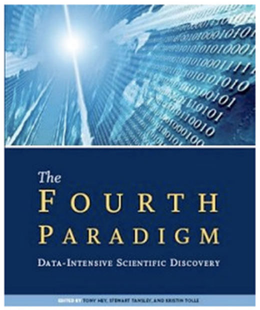
- **Computational Science**
  - Simulate analytical model
  - Validate model and makes predictions

- **Data Exploration Science**
  - **Data-driven science**
    Data captured by instruments
    or data generated by simulation
  - Processed by software
  - Placed in a database / files
  - Scientist(s) analyze(s) database / files
  - *Access crucial*

Fourth Paradigm —*Jim Gray's*

The Fourth Paradigm: Data-Intensive Scientific Discovery

Presenting the first broad look at the rapidly emerging field of data-intensive science



Increasingly, scientific breakthroughs will be powered by advanced computing capabilities that help researchers manipulate and explore massive datasets.

The speed at which any given scientific discipline advances will depend on how well its researchers collaborate with one another, and with technologists, in areas of eScience such as databases, workflow management, visualization, and cloud computing technologies.

In *The Fourth Paradigm: Data-Intensive Scientific Discovery*, the collection of essays expands on the vision of pioneering computer scientist Jim Gray for a new, fourth paradigm of discovery based on data-intensive science and offers insights into how it can be fully realized.

Download
- Full text
- By chapt

Related Re
- Microsoft
- eScience

Things are similar,
this makes science possible.

Things are different,
this makes science necessary.

Things are similar,
this makes Statistics *possible.*

Things are different,
this makes Statistics *necessary.*

BIG Statistics

- Business Statistics
- Industrial Statistics
- Governmental/Official Statistics (Official Statistics)

Data Science

BIG Data
High Dimensional Data

The Problem is not only the data is big, but what we're looking for is small… DeVaeux

Time Series Analysis

Time Series Analysis

- Time Series Analysis (Univariate)
  - Linear, ARIMA
  - Non-linear
- Time Series Analysis (Multi-variate)
- ARCH & GARCH Model
  - AutoRegressive Conditional Heteroscedasticity
  - GARCH, GARCH-M, Expo-GARCH etc
- Long-Memory, high-frequency data
- Ito's Lemma and Black-Scholes Differential Equation
- Financial Time Series
- Value at Risk

Time Series101: Univariate Time Series

135, 143, 127, 129, …, 172 What's next?

$y_1, y_2, \ldots, y_T$ What's $\hat{y}_{T+1}$?

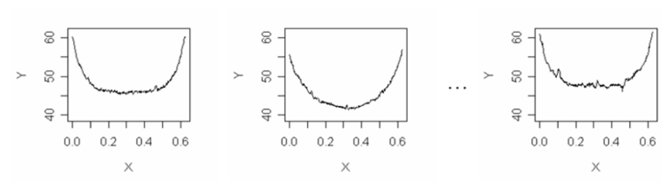
Model Building and Forecasting (short/long term)
Monitoring (Quality Assurance)
Change Point Problem

Time Series201: Multivariate Time Series

$$\begin{pmatrix} 5 \\ 143 \\ 85 \\ 46 \end{pmatrix}, \begin{pmatrix} 12 \\ 174 \\ 77 \\ 39 \end{pmatrix}, \ldots, \begin{pmatrix} 16 \\ 191 \\ 81 \\ 41 \end{pmatrix} \qquad \text{What's next?}$$

$$\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_T \qquad \text{What's } \hat{\vec{y}}_{T+1}?$$

Time Series301: Functional Data (Profile)

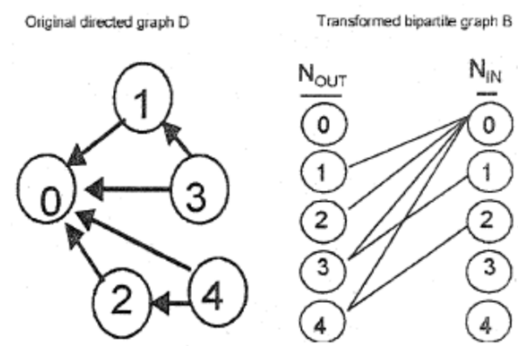


$$f_1, f_2, \ldots, f_T \qquad \text{What's } \hat{f}_{T+1}?$$

Time Series 401: Graphic/Network
(Link Prediction, Communication Network)

$$G_1, G_2, \ldots, G_T \qquad \text{What's } \hat{G}_{T+1}?$$

More: Directional Graphs



The Annals of Applied Statistics
2012, Vol. 6, No. 1, 409–427
DOI: 10.1214/11-AOAS501

PROFILE CONTROL CHARTS BASED ON NONPARAMETRIC L-1 REGRESSION METHODS[1]

BY YING WEI[2], ZHIBIAO ZHAO[3] AND DENNIS K. J. LIN

Columbia University, Pennsylvania State University and Pennsylvania State University

Classical statistical process control often relies on univariate characteristics. In many contemporary applications, however, the quality of products must be characterized by some functional relation between a response variable and its explanatory variables. Monitoring such functional profiles has been a rapidly growing field due to increasing demands. This paper develops a novel nonparametric L-1 location-scale model to screen the shapes of profiles. The model is built on three basic elements: location shifts, local shape distortions, and overall shape deviations, which are quantified by three individual metrics. The proposed approach is applied to the previously analyzed vertical density profile data, leading to some interesting insights.

INFORMS Journal on Computing
Vol. 21, No. 2, Spring 2009, pp. 286–303
ISSN 1091-9856 | EISSN 1526-5528 | 09 | 2102 | 0286

informs.

DOI 10.1287/ijoc.1080.0292

The Time-Series Link Prediction Problem with Applications in Communication Surveillance

Zan Huang, Dennis K. J. Lin
Department of Supply Chain and Information Systems, Smeal College of Business, Pennsylvania State University, University Park, Pennsylvania 16802 {zanhuang@psu.edu, dennislin@psu.edu}

The ability to predict linkages among data objects is central to many data mining tasks, such as product recommendation and social network analysis. Substantial literature has been devoted to the *link prediction* problem either as an implicitly embedded problem in specific applications or as a generic data mining task. This literature has mostly adopted a *static graph* representation where a snapshot of the network is analyzed to predict hidden or future links. However, this representation is only appropriate to investigate whether a certain link will ever occur and does not apply to many applications for which the prediction of the *repeated* link occurrences are of primary interest (e.g., communication network surveillance). In this paper, we introduce the *time-series link prediction problem*, taking into consideration temporal evolutions of link occurrences to predict link occurrence probabilities at a particular time. Using Enron e-mail data and high-energy particle physics literature coauthorship data, we have demonstrated that time-series models of single-link occurrences achieve comparable link prediction performance with commonly used static graph link prediction algorithms. Furthermore, a combination of static graph link prediction algorithms and time-series models produced significantly better predictions over static graph link prediction methods, demonstrating the great potential of integrated methods that exploit both interlink structural dependencies and intralink temporal dependencies.

Statistical Process Control for Monitoring Nonlinear Profiles: A Six Sigma Project on Curing Process

Shing I Chang[1], Tzong-Ru Tsai[2], Dennis K.J. Lin[3], Shih-Hsiung Chou[1] & Yu-Siang Lin[4]

[1]Quality Engineering Laboratory, Department of Industrial and Manufacturing Systems Engineering, Kansas State University, USA
[2]Department of Statistics, Tamkang University, Danshui District, New Taipei City 25137 Taiwan
[3]Department of Statistics, Pennsylvania State University, USA
[4]Department of Industrial Management, National Taiwan University of Science and Technology, Taipei, Taiwan

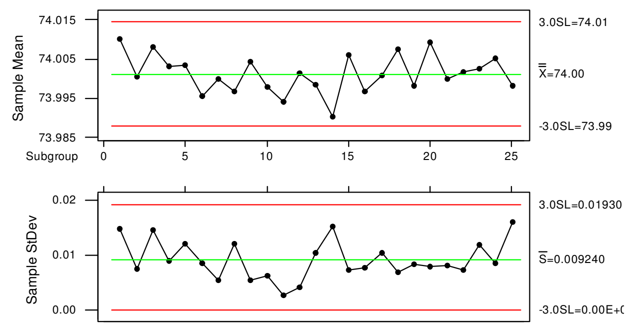
Quality Engineering

Statistical Process Control: Control Chart

What (sufficient statistics) to monitor
and its elated distribution
Statistical Hypothesis Testing in action!

Shewhart Control Charts
(Monitoring the Process)

Xbar and S Chart for Piston Ring Example



What to plot?
How to read?

Phase I

Phase II

Phase I of SPC

- Understand the variation in a process over time.
- Evaluate the process stability.
- Model the in-control process performance.
- Evaluated by Probability of Signal (POS)
- Focus on Robustness

Phase II of SPC

- Monitoring the process using on-line data
- Quickly detect parameter changes from the in-control parameter values
- Evaluated by Run Length (Average Run Length)
- Focus on sensitivity

How to compare two (Phase II) control charts?

Run Length Distribution

Run Length Distribution

- Run Length (T)
  - The period that control chart first signals
- Run Length Distribution f(T)
  - Distribution of Run Length
  - Theoretical & Empirical
- ARL—Average Run Length E(T)
  - How long, in the sense of average, we could expect the chart to detect signal

ARL has been used as gold standard

Marit Schoonhoven and Ronald J.M.M. Does, (2012), A Robust Standard Deviation Control Chart, **Technometrics**

Philippe Castagliola, Giovanni Celano and Stelios Psarakis, SV, (2011), Monitoring the coefficient of Variation Using EWMA Charts, **Journal of Quality Technology**;

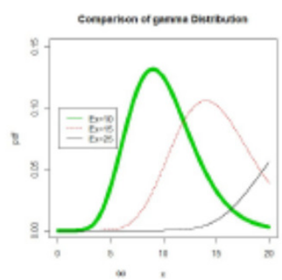
DAVIS, ROBERT B. and WOODALL, WILLIAM H., (2002) Evaluating and Improving the Synthetic Control Chart, **Journal of Quality Technology**;

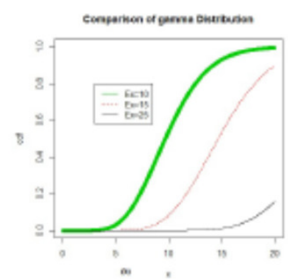
....

N. L. Johnson, (1961), Simple Theoretical Approach to Cumulative Sum Control Chart, **Journal of the American Statistical Association**

G. A. Barnard, (1959), Control Charts and Stochastic Processe, **Journal of the Royal**

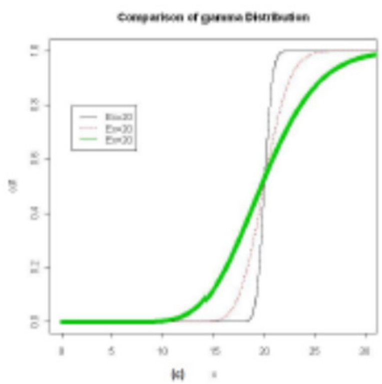
How informative is ARL?



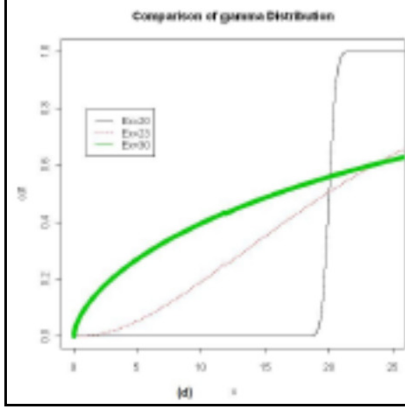


ARL=10 for the bold line;
ARL=15 for the dashed line;
ARL=25 for the solid line.

ARL=20 for all

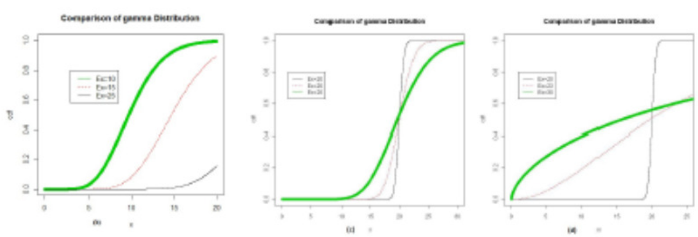


ARL may be misleading…



ARL=20 for the solid line;
ARL=23 for the dashed line;
ARL=30 for the bold line.

ARL is not so informative.
Is there a better criterion possible???



CRPS: Continuous Ranked Probability Score

$$CRPS(F,x) = \int_{-\infty}^{\infty} (F(y) - H(y-x))^2 dy$$

Where F denotes a cumulative distribution function, x is the target value, and H(y-x) denotes the Heaviside function and takes the value 0 when $y < x$ and the value 1 otherwise.

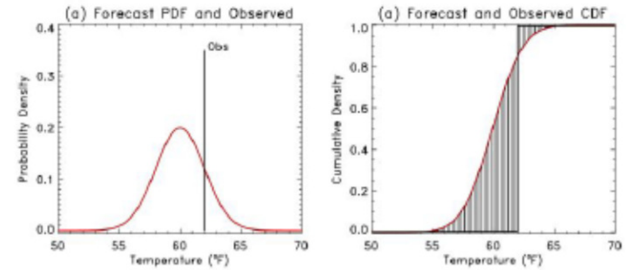
CRPS is popular in probability forecasts

Tilmann Gneiting and Adrian E. Raftery, (2007) Strictly Proper Scoring Rules, Prediction, and Estimation, **Journal of the American Statistical Association**;

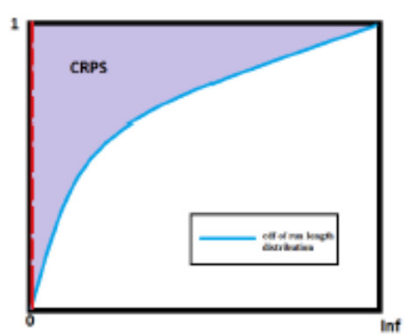
B. F. Gneiting, Tilmann and A. E. Raftery, (2007) Probabilistic forecasts, calibration and sharpness. **Journal of Royal Statistical Society, Ser. B**;

Hans Hersbach, (2000) Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems. **Weather and Forecasting**

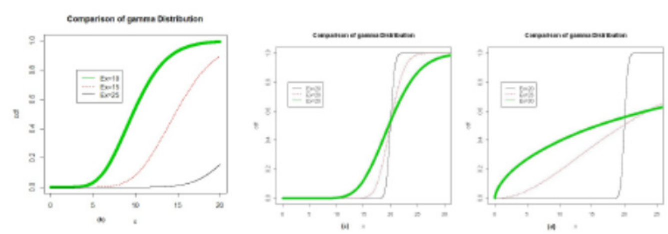
CRPS: Geometrical explanation



Applying CRPS to Run Length Distribution



CRPS vs ARL



Case	ARL	CRPS	Case	ARL	CRPS	Case	ARL	CRPS
I	10	8.238029	II	20	19.64322	III	20	19.74770
	15	12.83303		20	18.87303		23	14.89295
	25	22.19312		20	17.49259		30	11.97875

General Performances of CRPS

- Example 1: ideal parameters for EWMA chart
- Example 2: Control charts for standard deviation
- Example 3: EWMA-γ & EWMA-γ^2 Charts

Example 1: EWMA Chart

$$EWMA_t = (1-\lambda)EWMA_{t-1} + \lambda y_t \qquad 0 \le \lambda \le 1$$

where y_t is the sample mean at time t, and λ is a smoothing constant.

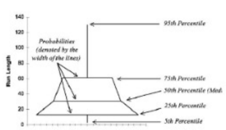
The control limit constant K is determined to yield the desirable in-control-ARL for a given λ.

Objective:
Choose the best combination of (λ,K).

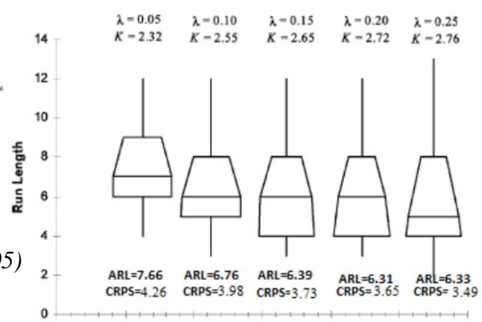
Crowder (1989) selected (0.20, 2.72) based on ARL.

PD-Plot

$$CRPS(EWMA) = \int_{-\infty}^{\infty}\left(\sum_{i=1}^{n}\int_{LCL}^{UCL} Pr(i-1,y) f\left(\frac{y-(1-\lambda)\mu}{\lambda}\right)dy - H(y-0)\right)^2 dy$$



PD-Plot



Radson & Boyd (2005) selected (0.25, 2.76), based on PD Plot.
This is consistent with CRPS!

Example 2: Control Charts for standard deviation: The comparisons via ARL and CRPS

— *Schoonhoven and Does (2012)*

		λ=0.6		λ=1.0		λ=1.2	
n	Chart	ARL	CRPS	ARL	CRPS	ARL	CRPS
5	$\check{S}$	131	65.24929	378	188.2596	69.5	33.49939
	D7	135	67.24969	373	186.2499	69.6	33.5494
	R^S	132	65.74968	369	184.2499	68.9	33.19939
	MD^S	135	67.24970	375	187.2499	69.7	33.5994
	MD^i	143	71.24972	371	185.2499	70.0	33.7494
	$MD^{i,S}$	143	71.24971	371	185.2499	69.9	33.6994
9	$\check{S}$	28.4	12.94851	371	185.2499	43.6	21.54903
	D7	29.5	13.49856	373	186.2499	43.9	21.69904
	R^S	29.2	13.34855	392	195.7499	43.9	21.69904
	MD^S	29.1	13.29854	369	184.2499	43.9	21.69904
	MD^i	29.8	13.64858	368	183.7499	44.7	22.09906
	$MD^{i,S}$	29.9	13.69853	366	182.7499	44.4	21.94905

The comparison results are consistent via ARL and CRPS. Why?

The Run Length Distribution here is believed to be Geometric Distribution —only one parameter involved!

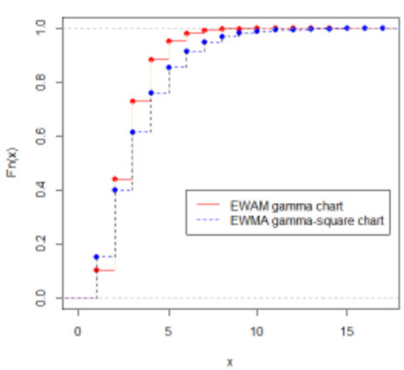
Example 3: EWMA-γ & EWMA-γ^2 Charts

Castagliola et al. (2011, JQT)

τ	Criterion	$\gamma_0 = 0.05$		$\gamma_0 = 0.10$		$\gamma_0 = 0.15$		$\gamma_0 = 0.20$	
	n=7	EWMA-γ^2	EWMA-γ	EWMA-γ^2	EWMA-γ	EWMA-γ^2	EWMA-γ	EWMA-γ^2	EWMA-γ
1.25	ARL	11.3	11.5	11.4	11.6	11.7	11.8	12.0	12.1
	CRPS	6.432	6.358	6.555	6.455	6.768	6.656	6.964	6.940
1.5	ARL	4.3	4.3	4.3	4.4	4.4	4.5	4.6	4.6
	CRPS	1.789	1.546	1.854	1.659	1.929	1.758	2.045	1.910
	n=10								
1.25	ARL	8.4	8.6	8.5	8.7	8.7	8.8	9.0	9.0
	CRPS	4.645	4.051	4.761	4.196	4.835	4.385	5.024	4.628
1.5	ARL	3.2	3.2	3.2	3.3	3.3	3.3	3.4	3.5
	CRPS	1.147	.793	1.186	0.831	1.277	0.982	1.301	1.046

The comparison results are different via ARL and CRPS. Which conclusion is more reliable???

Case (n, τ, γ_0)=(10, 1.5, 0.2)



τ	Criterion	$\gamma_0 = 0.20$	
	n=7	EWMA-γ^2	EWMA-γ
1.25	ARL	12.0	12.1
	CRPS	6.964	6.940
1.5	ARL	4.6	4.6
	CRPS	2.045	1.910
	n=10		
1.25	ARL	9.0	9.0
	CRPS	5.024	4.628
1.5	ARL	3.4	3.5
	CRPS	1.301	1.046

More About CRPS

$$CRPS(F,x) = \int_{-\infty}^{\infty} (F(y) - H(y-x))^2 dy$$

- Nau (9185): For any distant measure d who satisfies triangle inequality, we have

$$CRPS(F_1,x) \geq CRPS(F_2,x) \quad \textit{if and only if} \quad d(F_1,x) \geq d(F_2,x)$$

smaller CRPS → closer distance

- CRPS can also be represented as

$$CRPS(F,x) = E_F|Y-x| - \frac{1}{2}E_F|Y-Y'|$$

For Normal(μ,σ^2), this is

$$CRPS(F,x) = \sigma\{\frac{x-\mu}{\sigma}[2\Phi(\frac{x-\mu}{\sigma}) - 1] + 2\varphi(\frac{x-\mu}{\sigma}) - \frac{1}{\sqrt{\pi}}\}$$

More About CRPS

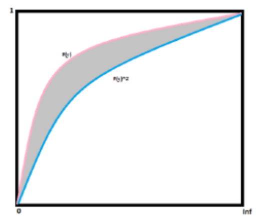
$CRPS(F,x)=\int_{-\infty}^{\infty}(F(y)-H(y-x))^2dy$

$$CRPS(F,x)=E_F|Y-x|-\frac{1}{2}E_F|Y-Y'|$$

For comparing Run Length Distribution, take x=0

Then
$E_F|Y-x|=E_F|Y|=E_F(Y)=ARL$
and

$\frac{1}{2}E_F|Y-Y'|=\int_{-\infty}^{\infty}F(y)(1-F(y))dy$
is also known as Gini Mean Difference.



CRPS = ARL + GMD

CRPS: Empirical Properties

Suppose F_n is the empirical function of F…

$$CRPS(F_n,x)=\int_{-\infty}^{\infty}(F_n(y)-H(y-x))^2dy\to\int_{-\infty}^{\infty}B^2(S(y))dy$$

where $B(u),(0\le u\le 1)$ is the classical Brownian bridge,

$$\sqrt{n}(F_n(y)-F(y))\xrightarrow{d}N(0,F(y)[1-F(y)])$$

(1) $CRPS(F_n,x)$ is an asymptotic unbiased estimator of $CRPS(F,x)$, *and*

(2) $CRPS(F_n,x)$ is strongly consistent to $CRPS(F,x)$.

Advantages of using CRPS

- Computationally straightforward
- Theoretical Support
  - CRPS=ARL + GMD (Gini Mean Difference)
- Empirical performance

- Paper submitted to *Technometrics*

Another Look at Run Length Distributions

Wei Wang
Department of Statistics, Penn State University
and
Dennis K.J. Lin
Department of Statistics, Penn State University

November 28, 2012

Direct Comparison Approach

Comparison without the exact Run Length Distribution

Comparison Algorithm

- Step 1: Construct the Control limits for all charts.
- Step 2: Simulate the out-of-control signal,
  - use all charts to detect the signal simultaneously.
  - Record the run length of first out of control signal for each chart.
- Step 3: Repeat Step 2 for sufficient long times (for example, 10,000 times).
- Step 4: Summarize the pairwise propositions of which chart detect the out of control signal earlier.

Three Popular Control Charts for Detecting Mean-Shift

$$N(\mu_1,\sigma^2) \rightarrow N(\mu_2,\sigma^2)$$

- X-bar Chart
- CUSUM Chart
- EWMA Chart

X-bar Chart

$$UCL = \mu + K\frac{\sigma}{\sqrt{n}}$$

$$Center\ line = \mu$$

$$LCL = \mu - K\frac{\sigma}{\sqrt{n}}$$

where K=2.88 for in-control ARL is 250

CUSUM Chart

$$CUSUM_i^+ = \max[0, x_i - (\mu + k) + CUSUM_{i-1}^+]$$
$$CUSUM_i^- = \max[0, (\mu + k) - x_i + CUSUM_{i-1}^-]$$
$$CUSUM_0^- = CUSUM_0^+ = 0$$

with the control limits

$$UCL = H$$
$$LCL = -H$$

where $H = 4.5\sigma$ and $k = 0.5\sigma$ for in-control ARL=250.

EWMA Chart

$$EWMA_i = (1-\lambda)EWMA_{i-1} + \lambda x_i$$

with the control limits

$$UCL = \mu + K'\sqrt{\frac{\lambda}{2-\lambda}}\frac{\sigma}{\sqrt{n}}$$
$$LCL = \mu - K'\sqrt{\frac{\lambda}{2-\lambda}}\frac{\sigma}{\sqrt{n}}$$

where $\lambda = 0.15$ and $K' = 2.65$ for in-control ARL=250.

Initial Results: Zero-State

Delta	$P(EWMA < Xbar)$	$P(CUSUM < Xbar)$	$P(EWMA < CUSUM)$
0.5	0.7829 (0.1007) 0.1164	0.8182 (0.0695) 0.1123	0.0733 (0.5058) 0.4209
1	0.7489 (0.1389) 0.1122	0.773 (0.1281) 0.1137	0.0409 (0.6546) 0.3045
1.5	0.6036 (0.2086) 0.1878	0.6219 (0.1908) 0.1873	0.0175 (0.766) 0.2165
2	0.4086 (0.2865) 0.3049	0.4279 (0.2768) 0.2953	0.0055 (0.8407) 0.1538
3	0.0839 (0.3272) 0.5889	0.0941 (0.338) 0.5679	**6e-04 (0.9194) 0.08**
4	0.0034 (0.2694) 0.7272	0.0057 (0.3263) 0.668	**0 (0.931) 0.069**

The values(left) given in the tables are the proportion of clear win, the values(right) given in the tables are the proportion of clear lose, and the proportion of tie is given in the parentheses

Delta	EWMA vs Xbar	CUSUM vs Xbar	EWMA vs CUSUM
0.5	EWMA	CUSUM	CUSUM
1	EWMA	CUSUM	CUSUM
1.5	EWMA	CUSUM	CUSUM
2	EWMA	CUSUM	CUSUM
3	Xbar	Xbar	★
4	Xbar	Xbar	★

Initial Results: Steady State

Delta	$P(EWMA < Xbar)$	$P(CUSUM < Xbar)$	$P(EWMA < CUSUM)$
0.5	0.777(0.0993) 0.1237	0.8505 (0.0555) 0.094	0.063 (0.2907) 0.6463
1	0.7574 (0.1341) 0.1085	0.7986 (0.0967) 0.1047	0.0723 (0.4261) 0.5016
1.5	0.6223(0.2078) 0.1699	0.6672 (0.1644) 0.1684	0.0845(0.5276) 0.3879
2	0.4422 (0.2874) 0.2704	0.4863 (0.2513) 0.2624	0.0939 (0.6015) 0.3046
3	0.1198 (0.3564) 0.5238	0.1575(0.3338) 0.5087	0.0882(0.7019) 0.2099
4	0.0202(0.374) 0.6058	0.0495 (0.349) 0.6015	0.078 (0.7704) 0.1516

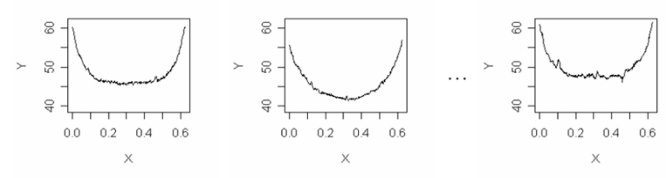
The proportions of Steady-State run length distributions comparisons

Delta	EWMA vs Xbar	CUSUM vs Xbar	EWMA vs CUSUM
0.5	EWMA	CUSUM	CUSUM
1	EWMA	CUSUM	CUSUM
1.5	EWMA	CUSUM	CUSUM
2	EWMA	CUSUM	CUSUM
3	Xbar	Xbar	CUSUM
4	Xbar	Xbar	★

Observations

- Zero-State and Steady-State show the similar results.
- When the mean-shift δ becomes large, the X-bar chart performs better, when δ is small, EWMA chart and CUSUM chart detect the out-of-control signal earlier.
- When δ becomes large, dependency exists between control charts due to the general large proportion of tie. We plan to extend such a comparison study to a general class of all control charts.
- Study the Pitman closeness for the simulative comparisons.

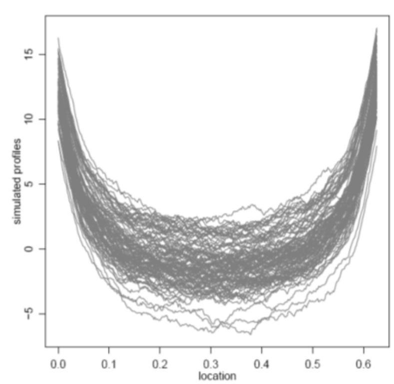
Monitoring Functional Data (Profile)



In general, there are more than one X.

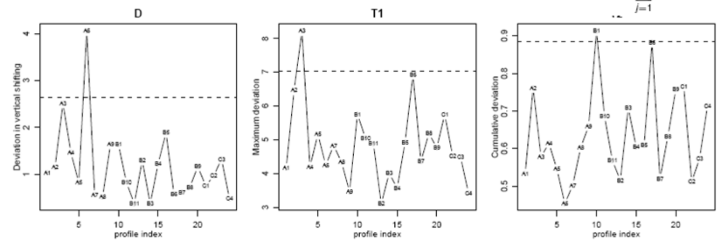
VDP Data

For vertical density close to the wood surface



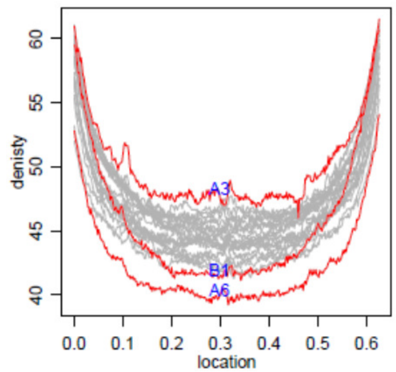
Upper Limits are 2.64, 7.02, and 0.88

$d_i = |\xi_i - \mu_\xi| / s_\xi$ $\qquad$ $T_m^{(1)} = \max_{1 \le j \le m} |\hat{e}_j|$ $\qquad$ $T_m^{(2)}(\lambda) = \sum_{j=1}^{m} |\hat{e}_j|^\lambda$

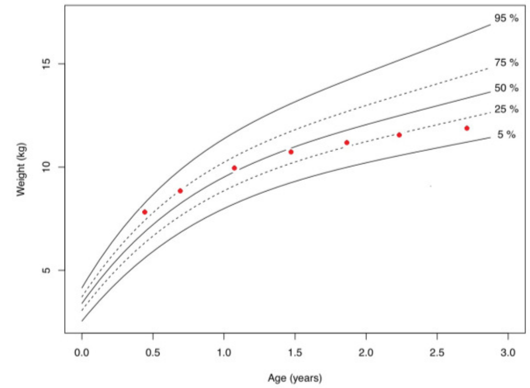


Profiles A6, A3 & B1 were identified!

VDP Data



Why Functional: Growth Curve Example



STILL QUESTION?

Send $500 to

- Dennis Lin

  University Distinguished Professor

  317 Thomas Building

  Department of Statistics

  Penn State University

- +1 814 865-0377 (phone)
- +1 814 863-7114 (fax)
- DennisLin@psu.edu

(Customer Satisfaction or your money back!)