Some aspects of target-setting in supply chain monitoring

Emmanuel Yashchin
IBM Research, Yorktown Heights, NY

JRC-2014, Seattle WA, June 2014

Outline

- Motivation
- Likelihood Ratio based monitoring
- Target setting based on engineering specifications
- Business considerations
- Date-driven target-setting
- Conclusions

Special thanks to IBM colleagues: R. Baseman, A. Civil, W. Hoffman, D. Jensen, J. Komatsu, S. Ruegsegger, J. Wargo, P. Zulpa,

Selected IBM activities in the SPM area

- Semiconductors
- Pension fund
- Supply chain
- Warranty data

Note: This job ran further analysis using **APCS** (Advanced Process Control System) from IBM Yorktown Watson Research Labs.

APCS found 1673 alarms from 31437 total charts. (Larger severity is worse.)

parmdesc	oper	tool	days ago	severity	link
Std_Min_N_PSP_Ion/Wdes_merge_H	RIEOffSetSpc51P_1	FJ02:PM1	8	57.506	view
Std_Min_N_PSP_Ion/Wdes_merge_H	CMPCuSiCOHM1P_1	ED11:LPOL	5	43.149	view
Std_Min_N_PSP_Ion/Wdes_merge_H	CMPSTIFA0RXP_1	EF01:Platen1	27	41.986	view
Std_Min_N_PSP_Ion/Wdes_merge_H	RTPSDActiveP_1	CBD1:ChC	2	41.027	view
Std_Min_N_PSP_Ion/Wdes_merge_H	WETHFBEOLM1P_1	DWD1:PMB	5	39.79	view
Std_Min_N_PSP_Ion/Wdes_merge_H	WETAEROSOLM1P_1	DAD1:Ch3	3	39.717	view
Std_Min_N_PSP_Ion/Wdes_merge_H	RIEDFS/LIMDSP_1	FSD1:PM3	3	38.375	view

Tips:You can view this table in its own new window if desired.

Here is some documentation on APCS.

Early Detection Tool Dashboard

Date of Analysis Jan 13, 2004 By Ship Vintage Latin America

	Consumer	Desktop	Mobile	PWS	Server
AC/DC Converters	0/13	0/165		0/41	0/124
CD/DVD	0/80	0/548	0/297	0/118	0/239
Flat Panel Displays			0/146		
Hard Disk Drives	0/38	8/382	0/410	1/55	0/165
Keyboards			0/250		
Rechargable Batteries			0/177		
System Planar	0/33	3/329	2/325	0/60	0/161

Part Number	Code Name	Description	Step Name	Trend Stoplight	AVT Stoplight	QEWS Stoplight	Drill to Cube
000002000563	HMS	HMS	BLFO				
000002002626	GUAYIN_MAOKONG	GUAYIN_MAOKONG	BLFO				
000002002929	YILAN PLANAR	YILAN PLANAR	BLFO				
000002003346	YUNGAN PLANAR	YUNGAN PLANAR	BLFO				NO CUBE AVAILABLE
000002000882	BEITOU	BEITOU	BLFO				
000002000900	YUSHAN	YUSHAN	BLFO				
000002002928	HARUHI	HARUHI	BLFO				
000002000478	MEIYA	MEIYA	BLFO				
000002000035	PIRANHA	PIRANHA	BLFO				

Framework and Approach

1. Data: Massive, sometimes high intensity

Typically unstable, possible outliers

2. Objectives

a. Alarms Believable and operationally (not statistically) significant
b. High statistical power
c. Detect all types of changes of interest
d. Facilitate alarm prioritization
e. Designability
f. Facilitate post-alarm diagnostics (esp. graphical analysis)

3. Approach: Parametric (3-zone)

a. Decompose data stream into sequences of statistics of interest
b. Specify acceptable/unacceptable regions for monitored parameters
c. Specify false alarm rate / sensitivity requirements
d. Deploy schemes based on Likelihood Ratios (Cusum-Shewhart)

Target setting

Likelihood Ratio (LR) Approach:

$f_0(\bullet)$ - "good" process $\qquad$ $f_1(\bullet)$ - "bad" process

LR Scheme: Signal at time T if for some $m \geq 1$

$$\sum_{i=T-m+1}^{T} \ln[f_1(X_t)/f_0(X_t)] > A$$

Alternative formulation (Page's scheme): Set

$$s_0 = 0, \quad s_t = \max\left[s_{t-1} + \ln\frac{f_1(X_t)}{f_0(X_t)}, 0 \right], \quad t = 1, 2, \ldots$$

and signal when $s_t > A$

Example: Normal (Gaussian) Case:

$f_0(\bullet)$ - Normal (μ_0,σ) $f_1(\bullet)$ - Normal (μ_1,σ)

Then: Score = $\ln \dfrac{f_1(X_t)}{f_0(X_t)} = \dfrac{1}{\Delta}\left(X_t - \dfrac{\mu_0 + \mu_1}{2} \right)$

LR (Cusum) Upper Scheme: Set $k = \dfrac{\mu_0 + \mu_1}{2}$ and select $h \geq 0$.

Signal a process change *upwards* when $s_t > h$, where

$$s_0 = 0, \quad s_t = \max\left[s_{t-1} + (X_t - k), 0 \right], \quad t = 1, 2, \ldots$$

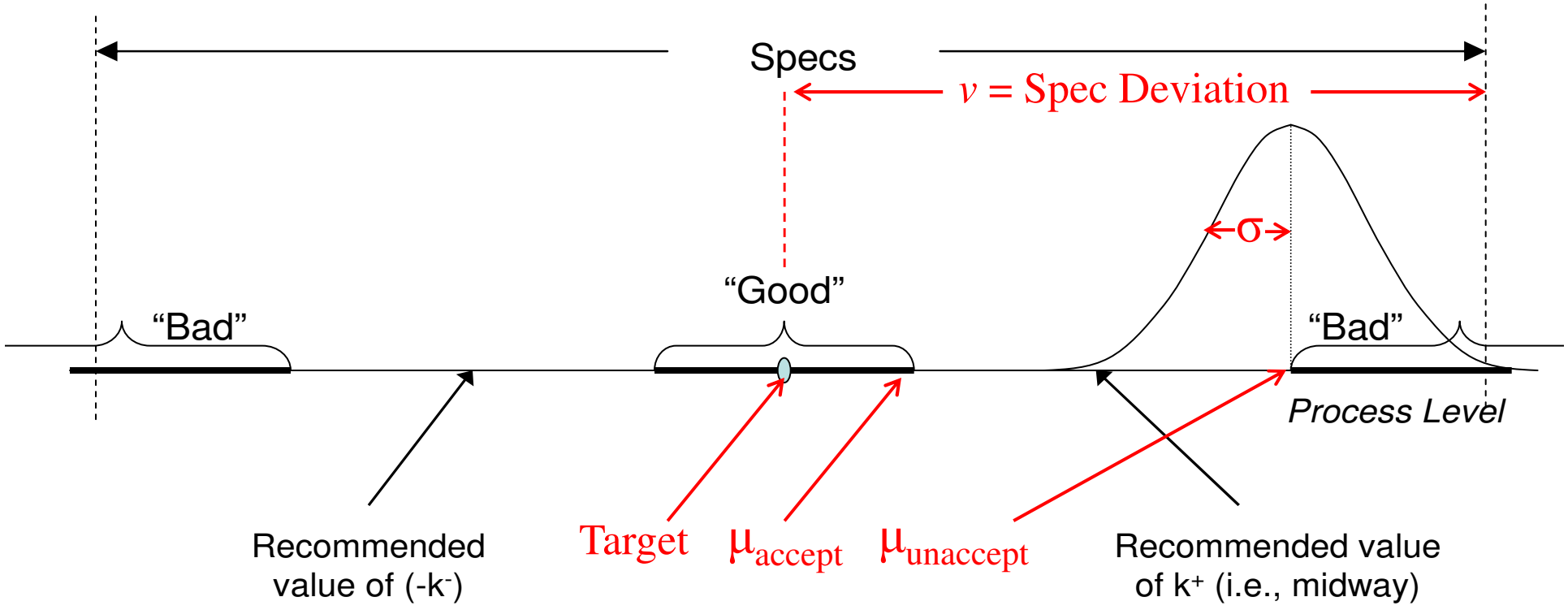
Note: Design of a scheme involves just *one "knob"*, i.e., *h.*

Target-setting when Specs are present

In this case, usually available: Target, Capability Indices

Need to find: acceptable/unacceptable mean levels μ_{accept}, $\mu_{unaccept}$

Process Window Configuration (2-sided case)



Target-setting process: Example

Performance of the monitoring system will *depend on level of input* that the user (real or virtual) is willing to provide

Complete or partially complete input: Target, σ, Acceptable/Unacceptable levels, type of control, False Alarm (FA) rate.

spcvar	target	pop. sigma	input format (Abs/Del)	Acceptable level	Unacceptable level	type	FA rate	Unacc. Sigma	spec deviation (for suggesting)	parm description
v001	0	10		-2	-10	2 sym. / 1 side	10	1.5	?:	DMIW ptile Std_N_
v002	18	.15		18.1	19	2 sym. / 1 side	5	1.5	?:	DMIW ptile PSRO_(
v003	5	01		5.1	5.2	2 sym.	10	1.5	?:	DMIW meas v003 M

Minimal input: type of control (A), FA rate and "spec deviation" (B)

A. B.

spcvar	target	pop. sigma	input format (Abs/Del)	Acceptable level	Unacceptable level	type	FA rate	Unacc. Sigma	spec deviation (for suggesting)	parm description
v001						2 sym. / 1 side	1000	1.5	?: 45	DMIW ptile Std_N_M1_Ioff . . . . .

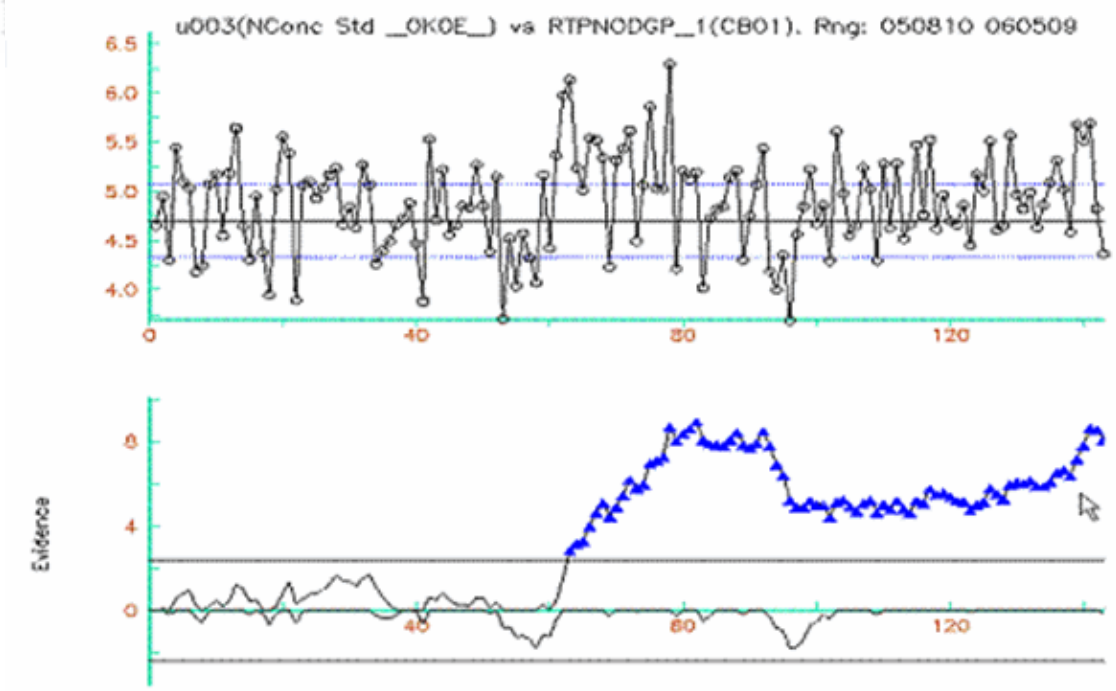
Follow-up run and tuning

Inspection of the *dashboard* and of *individual charts*

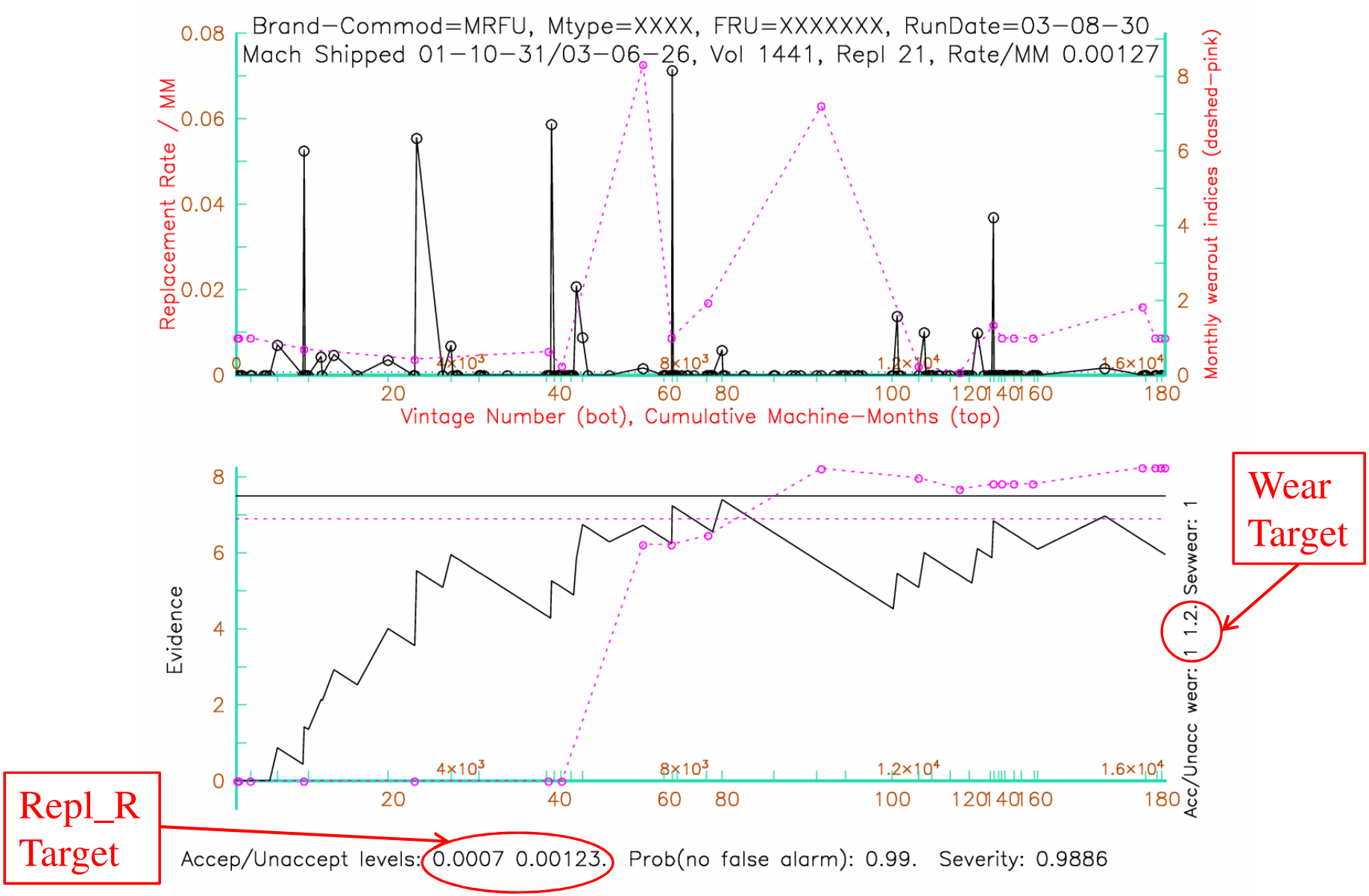
Note: This job ran further analysis using **APCS** (Advanced Process Control System) from IBM Yorktown Watson Research Labs.

APCS found 1673 alarms from 31437 total charts. (Larger severity is worse.)

parmdesc	oper	tool	days ago	severity	link
Std_Min_N_PSP_Ion/Wdes_merge_H	RIEOffSetSpcr51P_1	FJ02:PM1	8	57.506	view
Std_Min_N_PSP_Ion/Wdes_merge_H	CMPCuSiCOHM1P_1	ED11:LPOL	5	43.149	view
Std_Min_N_PSP_Ion/Wdes_merge_H	CMPSTIFA0RXP_1	EF01:Platen1	27	41.986	view
Std_Min_N_PSP_Ion/Wdes_merge_H	RTPSDActiveP_1	CBD1:ChC	2	41.027	view
Std_Min_N_PSP_Ion/Wdes_merge_H	WETHFBEOLM1P_1	DWD1:PMB	5	39.79	view
Std_Min_N_PSP_Ion/Wdes_merge_H	WETAEROSOLM1P_1	DAD1:Ch3	3	39.717	view
Std_Min_N_PSP_Ion/Wdes_merge_H	RIEDESCUMDSP_1	ESD1:PM3	3	38.375	view



Targets can be driven by service cost estimation (monitoring Replacement Rate & Wearout)



Target Setting when no Specs available

Example: Supply Chain parts flow process (three-silo stream)

Tests: Parts subject to various tests in every silo

a. Tested as individual parts or within sub-systems
b. Results: Pass/Fail (mostly)
c. Several (sometimes many) parts of similar type, e.g. cables, hard drives
d. Generation transitions typical

Targets: Driven by Quality Management Program

a. Business requirements
b. Historic performance
c. Costs
d. Targets needed *for every part/silo/test*

Basic approach: Weighted Cusum

- Denote $\{P_i, i = 1, 2, \ldots\}$ proportions of failed parts in successive weeks.
- Treat weekly number of units tested as weights $\{w_i, i = 1, 2, \ldots\}$ for $\{P_i\}$
- Monitored parameter: $\lambda = E(P_i)$
- *Establish acceptable & unacceptable regions* $\lambda_0 < \lambda_1$
- *Establish acceptable rate of false alarms*
- Apply Weighted Cusum (Page's) scheme to every relevant data set; flag this data set if out-of-control conditions are present

Repeated Page's scheme

Suppose that at time T we have data for N vintages

Define the set $\{S_i\ ,\ i = 1, 2, \ldots, N\}$ as follows:

$$S_0 = 0, \quad S_i = \max[0, \gamma S_{i-1} + w_i(P_i - k)],$$

where

$$k = (\lambda_1 - \lambda_0) / \ln(\lambda_1 / \lambda_0) \approx (\lambda_1 + \lambda_0) / 2, \quad \gamma \in [0.7, 1]$$

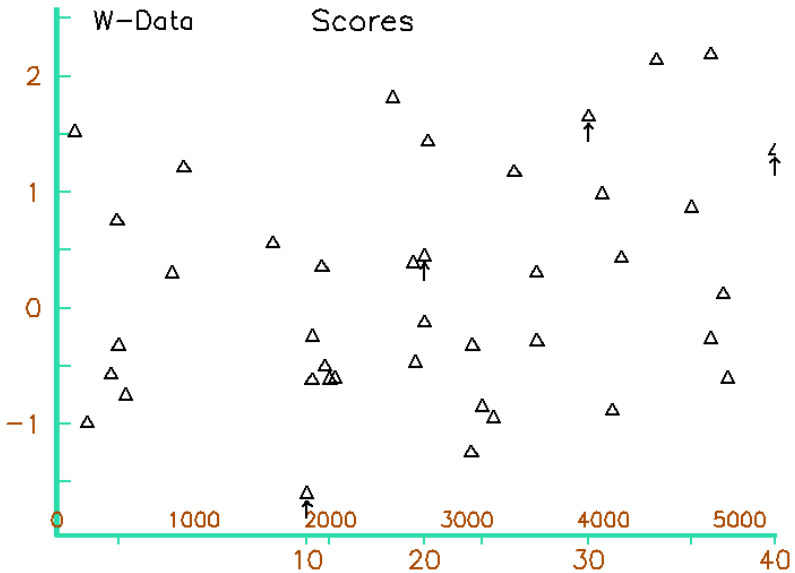
Define $S = max\ [S_1, S_2, \ldots, S_N]$;

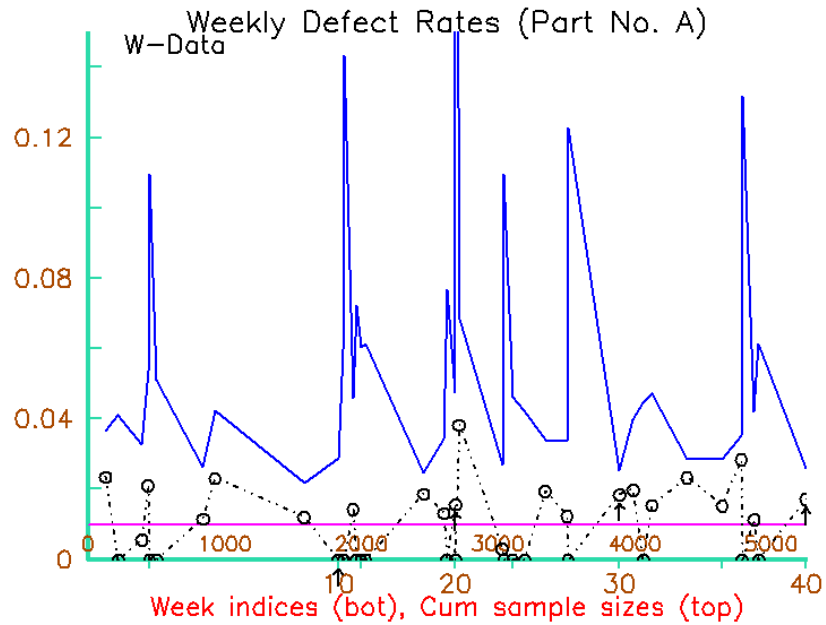
Flag the data set at time T if $S > h$, where h is chosen via:

$$\text{Prob}\{\ S > \text{h} \mid N, w_1, w_2, \ldots, w_N, \lambda = \lambda_0\} = \alpha_0 \ \ (\text{e.g.} = 0.01)$$

Supplemental tests: typically needed, e.g., Shewhart bounds

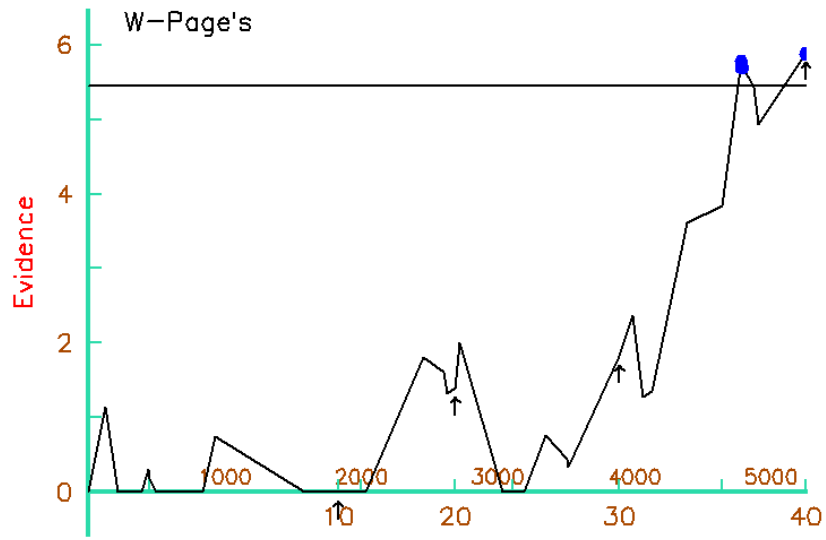
Weighted Page's analysis: example





Week, Defects, Sample Sizes

1	3	127	16	0	34	31	2	100
2	0	93	17	8	424	32	0	74
3	1	172	18	2	150	33	1	64
4	1	47	19	0	20	34	6	256
5	0	9	20	1	63	35	4	258
6	0	53	21	0	1	36	4	140
7	4	341	22	1	26	37	0	6
8	2	86	23	1	317	38	1	87
9	8	653	24	0	9	39	0	34
10	0	248	25	0	68	40	6	345
11	0	36	26	0	85			
12	0	5	27	3	153			
13	1	69	28	2	159			
14	0	23	29	0	7			
15	0	35	30	7	377			



Accept/Unaccept levels: 0.01 0.02. False alarm rate: 370

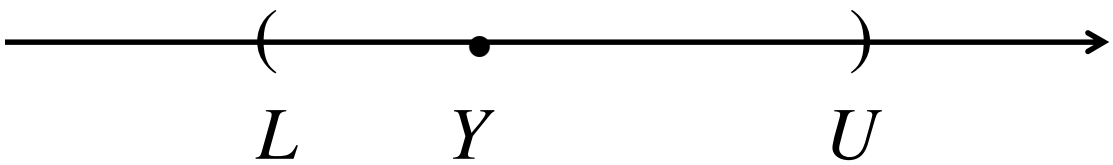
Threshold Accept level Shewhart limit

Target derivation process

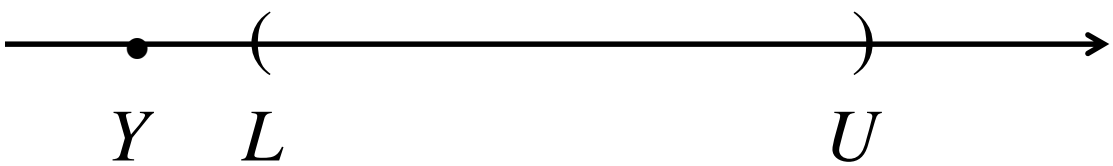
Target T for the fallout rate p of a given Part A is based on:

a. Robust (say, 90%) confidence bounds *(L,U)* for *p*

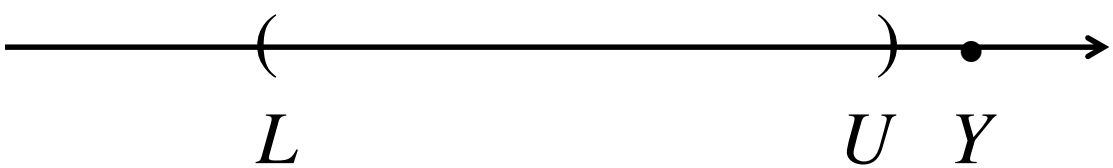
b. Yardstick *Y* based on Part A hierarchy (*Y*="peer" performance)

Case 1: 

Select $T=Y$

Case 2: 

Select $T=L$

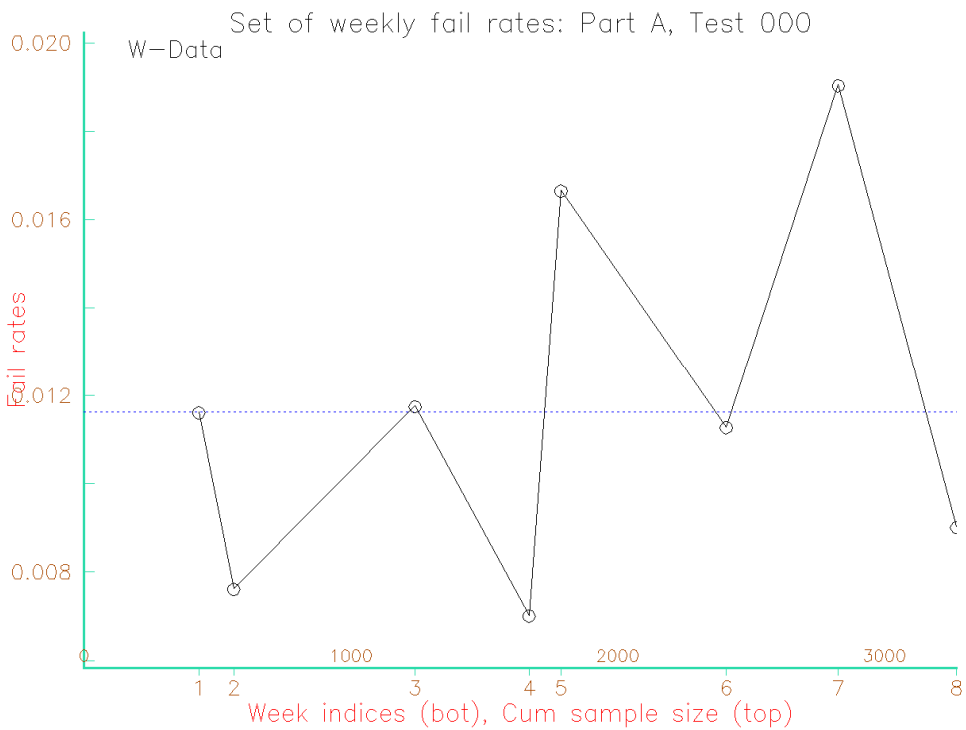
Case 3: 

Select $T=U$

This ensures that T is consistent with "capability"

Robust estimation for weighted data

Goal: *Obtain estimates and confidence bounds (L,U) for p (Trimmed Weights estimation).*



Given are:

a. Sequence $\{P_1, P_2, \ldots, P_N\}$ of estimates corresponding to sample sizes $\{n_1, n_2, \ldots n_N\}$.

b. Trimming proportion α

Denote:

$\{X_1, X_2, \ldots, X_N\}$ = sequence of fail counts corresponding to $\{P_i, n_i\}$

"Usual" weighted estimate $\hat{p}_w$

Index	1	2	3	4	5	6	7	8	Totals
X_i	5	1	8	3	2	7	8	4	38
N_i	430	131	680	427	120	620	420	443	3271
P_i	.0116	.0076	.0118	.0070	.0167	.0113	.0190	.0090	.0116

Robust estimation for weighted data (cont)

To obtain a Trimmed Weights Estimator:

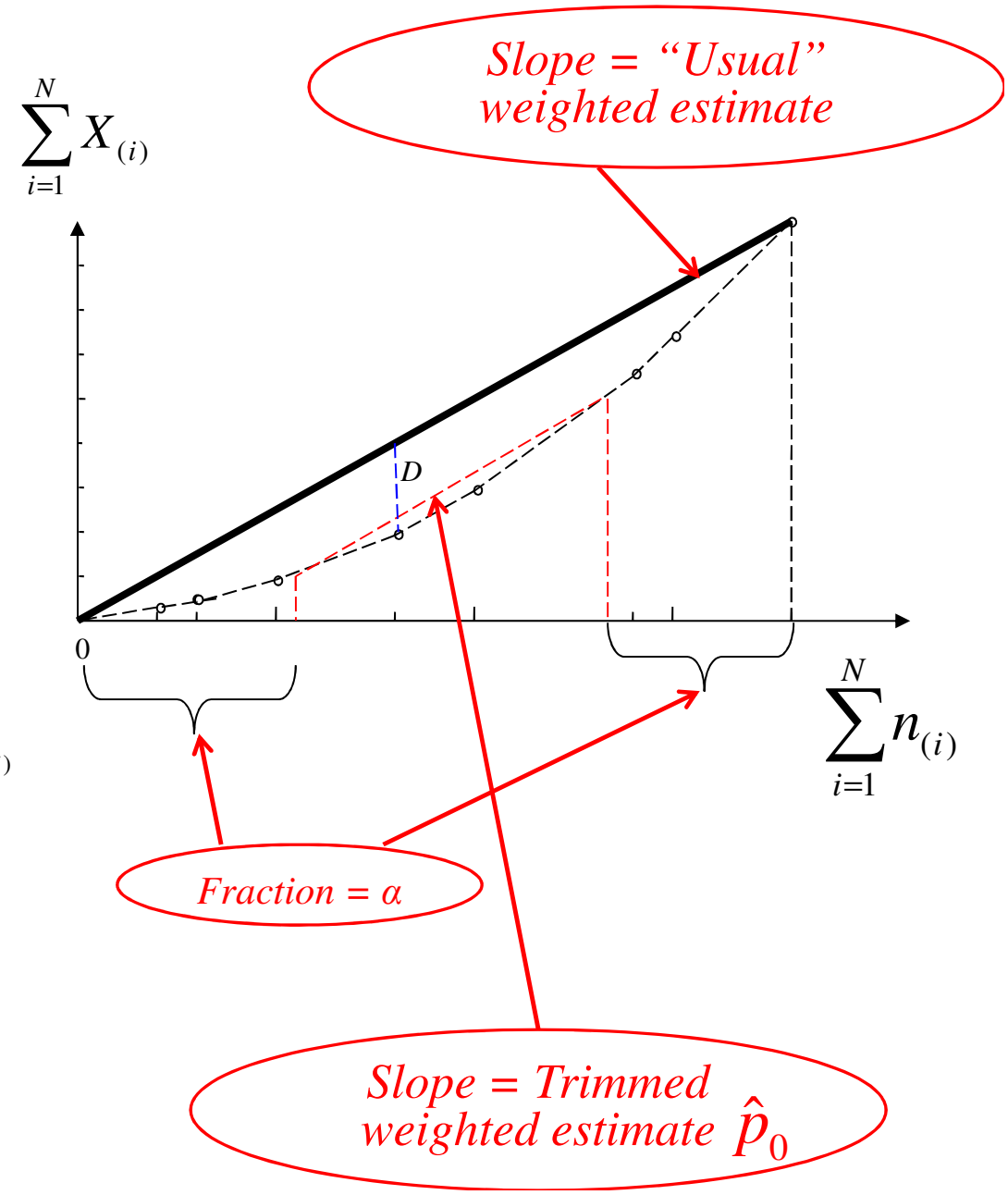
a. Sort $\{P_1, P_2, \ldots, P_N\}$ in increasing order => $\{P_{(1)}, P_{(2)}, \ldots, P_{(N)}\}$

b. Let $\{X_{(i)}, n_{(i)}\}$ be number of fails & sample sizes corresp. to $P_{(i)}$ (*note:* $P_{(i)} = X_{(i)}/n_{(i)}$)

c. Produce plot of cum $X_{(i)}$ vs cumulative $n_{(i)}$

d. Interpolate above plots at fractions α and (1- α) of $\sum_{i=1}^{N} n_{(i)}$

e. Set $\hat{p}_0$ = *Slope of interpolating line*



Confidence bounds

To obtain the (1-β)•100% ***bounds (L, U) for p:***

a) Set the effective overall sample size

$$n_{eff} = (1 - 2u\alpha)\sum_{i=1}^{N} n_i$$

b) Define the equivalent observed number of fails by

$$X_{eff} = n_{eff} \bullet \hat{p}_0$$

c) Let *F(x,a,b)* = Beta-cdf. Upper/Lower CB found by solving:

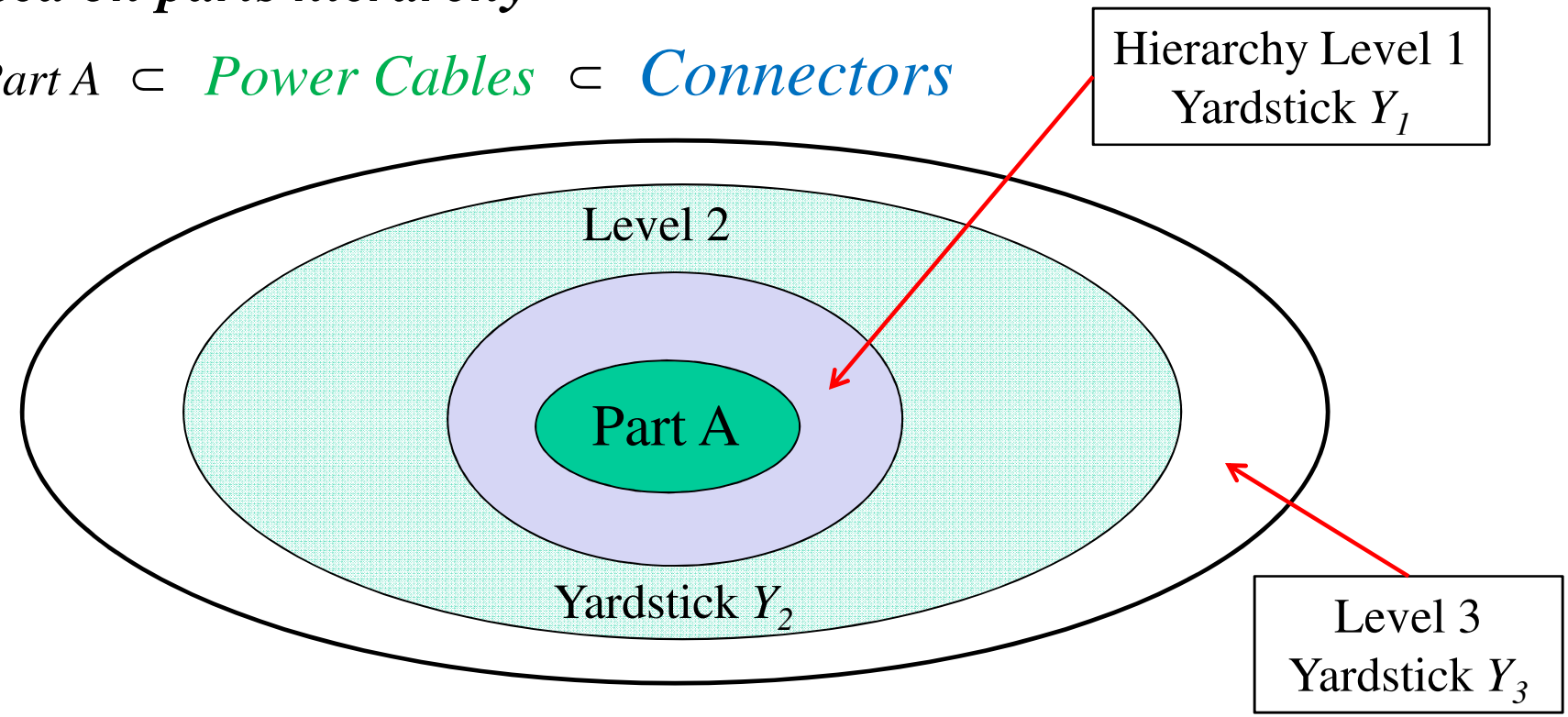
$$1 - \beta = F(x, X_{eff} + 1, n_{eff} - X_{eff})$$

$$\beta = F(x, X_{eff}, n_{eff} + 1 - X_{eff})$$

Constructing Yardstick Y

Based on parts hierarchy

E.g. *Part A* ⊂ *Power Cables* ⊂ *Connectors*



Sequential process: stop at level K for which coverage criteria are met

$$Y = a_1 Y_1 + a_2 Y_2 + \ldots + a_K Y_K$$

E.g. Step$_1$: a_1=.5; Step$_2$: a_1=.5, a_2=.3; Step$_3$: a_1=.5, a_2=.3, a_3=.2 (Sum=1)

Conclusions

1. Efficient target-setting essential for success of a system for monitoring massive data
2. Targets associated with parameters of statistical models and robust estimation are received well by the engineers in conjunction with likelihood-ratio based monitoring
3. Target-setting is a multi-objective process
4. Incorporate factors related to business needs (costs), engineering requirements (specs), peer performance (hierarchy), historic data features.
5. Multivariate extensions