CUSUM for Counts: Power Considerations and the Low Count Regime

by

Darwin J. Davis	*University of Delaware*
James M. Lucas	*J.M. Lucas and Associates*
Erwin M. Saniga	*University of Delaware*
Thomas P. McWilliams	*Drexel University*
Emma Pollard	*University of Delaware*

Presenter:

James M. Lucas JamesM.Lucas@Verizon.net

61st Annual Fall Technical Conference

October 5-6, 2017 - Philadelphia, PA

Abstract: Our Research Project

A CUSUM is an optimum detector of a change in distribution. For shift detection, a CUSUM will detect a specified shift faster than any other control procedure that has the same in-control ARL (false alarm frequency). A CUSUM for detecting shifts for normal variables in a "standard" situation, when the in-control ARL is within a factor of 3 of the ARL of a 3σ Shewhart Chart ($370/3 < ARL < 3x370$), can be designed to detect shifts larger than 0.5σ. In "data rich" situations larger in-control ARLs can be used and smaller sifts can be detected. For detecting a one σ shift a reference value ($k = 0.5$ or half of the standardized shift) and a decision interval value ($4 \leq h \leq 5$) will often be used. For counts the situation is more complicated. Only in the high count regime will the above guidelines work well because then the count distribution is well approximated by the normal distribution. In the medium and low count regimes, there is a different count distribution for each situation. This means that the tables needed to design the appropriate CUSUM are extensive and tables for all count situations have not been published. We have developed automatic design procedure for the various count distributions. These determine the appropriate parameter values for the specified count control situation and describe the characteristics of the specific recommended control procedure of interest.

Purpose:
To describe the usefulness of power for monitoring applications and to give a complete set of recommendations for count monitoring when count levels are low

Abstract:
In this talk we discuss the power needs of control schemes and give a set of recommendations for the low-count regime. We show that the power needed by monitoring applications is less than the power required by other applications. We show that an ARL Ratio ($ARL_{in\text{-}control}/ARL_{out\text{-}of\text{-}control}$) > 20 gives adequate power for monitoring applications.

The low count regime has the complication in that only large proportional shifts in level can be detected. We discuss the ability of monitoring procedures to detect order of magnitude (OOM) shifts, OOM/2 shifts and a doubling in the low-count regime. We show that it is important to consider not only the size of the shift but also the “data richness” so that the amount of data needed to detect the shift with high power will be obtained. For detecting doubling, we compare the rule of 50 used in clinical trials with a rule of 20 that works for monitoring applications. We also show that, when count levels are low, it is usually not feasible to detect improvement so only high-side control is generally recommended.

A CUSUM is an optimum detector of a change in distribution. For shift detection, a CUSUM will detect a specified shift faster than any other control procedure that has the same $ARL_{in\text{-}control}$ (false alarm frequency). Trying to detect too small a shift can lead to an optimum procedure with poor performance. Optimum does not always imply good.

Purpose:

- To completely describe control situations in the Low Count Regime (with $\mu < 0.04$)
- This description uses a Stoplight Model for Power and recognizes the appropriate Sampling Environment (giving an ARL_{goal})
- Sampling Environments include Standard, Data Rich and Big Data

Some Applications

- Manufacturing Processes
  - Counts of nonconformances
  - Low frequency Quality Problems
  - Tracking changes in purchase rates
- Occurrence of Adverse Events
  - Congenital malformations
  - Surgical Mortality by procedure
  - Flu victims
  - High School Football Deaths
  - Victims of violence

CUSUM for Counts: Overview

- 1) Introduction to CUSUM
- 2) Power Considerations
  - Stop Light Model
- 3) Sampling Environments
  - Sparse, Standard, Data Rich and Big Data Environments
    - Sparse Data - Short Production Runs
- 4) Regimes for Counted Data
  - Low, Intermediate & High Count
  - The Low Count Regime ($\mu < 0.04$)
- 5) Control in the Low Count Regime
  - Order of Magnitude (OOM), OOM/2 & Doubling Shifts
  - Detecting Improvement
- 6) More About Detecting Doubling
  - Rule of 50 & 20
- 7) Conclusions

1) Introduction to CUSUM

- CUSUM, based on a Likelihood Ratio test, is an optimal detector of a change in distribution (Moustakides, G.V. 1986)
- CUSUM Performance is often evaluated by the ARL

Calculation of k

- The "k" formula is based on a Sequential Probability Ratio Test (Wald's SPRT)

Binomial

$$k = \frac{n * \ln\left[\frac{1-p_0}{1-p_d}\right]}{\ln\left[\frac{1-p_0}{1-p_d}\right] - \ln\left[\frac{p_0}{p_d}\right]}$$

n = sample size

p_0 = in-control proportion defective

p_d = out-of-control proportion defective

Poisson

$$k = \frac{(\mu_d - \mu_a)}{\ln(\mu_d) - \ln(\mu_a)}$$

μ_a = mean in control

μ_d = mean out of control

CUSUM Implementation Formulas

High Side CUSUM and Low Side CUSUM Formulas

S_i = CUSUM at stage i

S_i = 0 for 0-state CUSUM; $S_i > 0$ (usually $S_i = h/2$) for FIR CUSUM

y_i = the observed count

k = reference value

h = decision interval

For the high side

$S_i = \text{Max} \{ 0, \; S_{i-1} + y_i - k_H \}$ signal if $S_i \geq h_H$

For the low side

$S_i = \text{Max} \{ 0, \; S_{i-1} + k_L - y_i \}$ signal if $S_i \geq h_L$

Implementation Example: Poisson CUSUM

(h = 9.3, k = 5.35)

i	Y_i	$Y_i - k$	S_i No FIR	S_i FIR
0			**0**	**4.65**
1	3	-2.35	0	2.3
2	7	1.65	1.65	3.95
3	2	-3.35	0	0.6
4	0	-5.35	0	0
5	2	-3.35	0	0
6	8	2.65	2.65	2.65
7	4	-1.35	1.3	1.3
8	0	-5.35	0	0
9	2	-3.35	0	0
10	3	-2.35	0	0
11	10	4.65	4.65	4.65
12	8	2.65	7.3	7.3
13	4	-1.35	5.95	5.95
14*	9	3.65	9.6	9.6
15*	11	5.65	15.25	15.25

*Out-of-Control Signal

*Out-of-Control Signal

For the high side

$$S_i = \text{Max}\{0,\ S_{i-1} + y_i - k_H\}$$

signal if $S_i \geq h_H$

While the implementation is simple, it cannot be improved upon because it is an optimum procedure (Moustakides, 1986)

Average Run Length (ARL) Performance

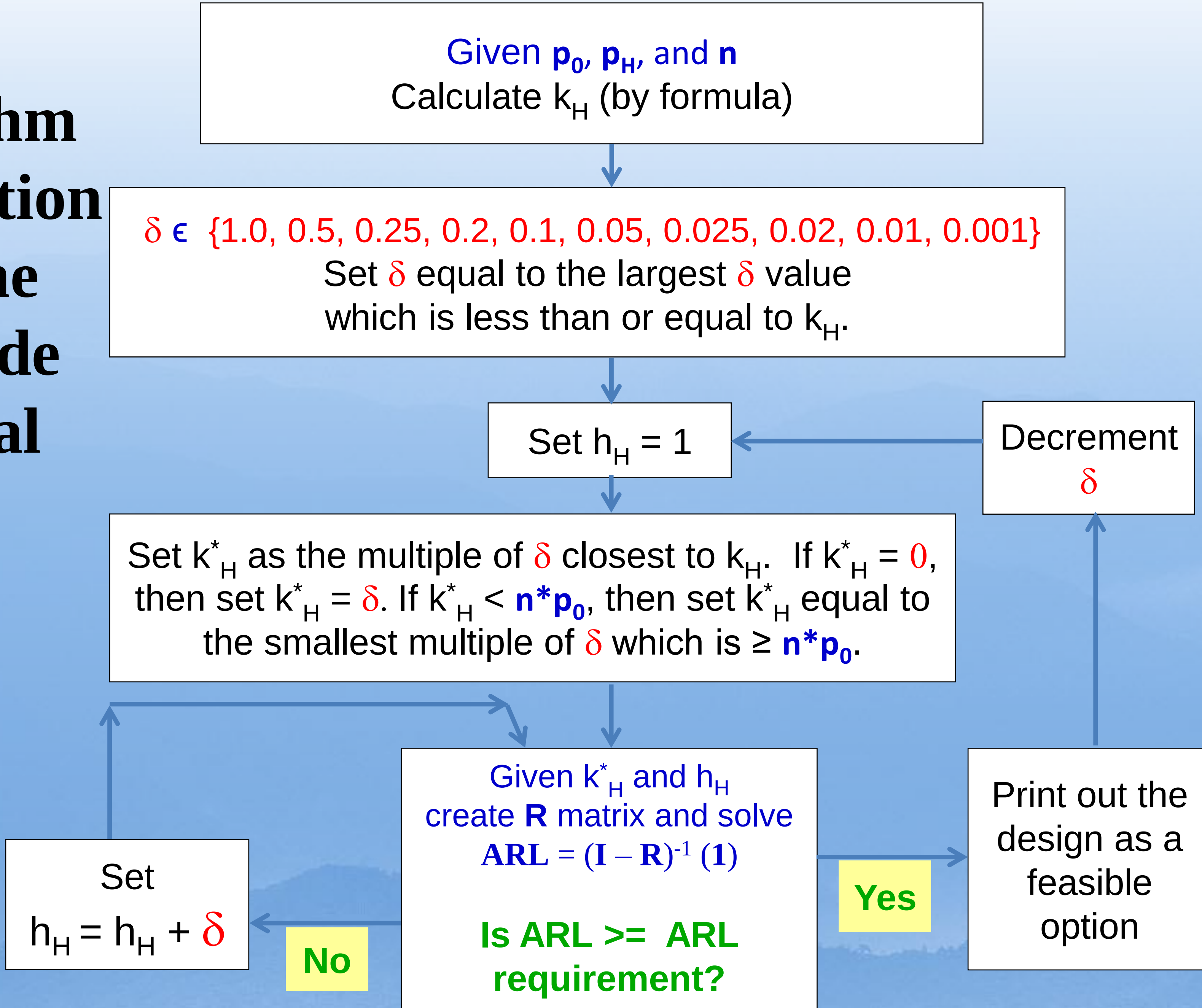
- ARL = Average Number of Samples before a Signal

- Desirable ARL Characteristics:
  - Process Out-of-Control (High Side) → Small ARL (quick detection)
  - Process In-Control → Large ARL (few false alarms)
  - Process Out-of-Control (Low Side) → Small ARL (quick detection)

CUSUM Design

- Calculate k using formula given previously
- Choose h to meet in-control ARL goals
  - Use the Search Algorithm given on next slide
  - Verify that parameters satisfy control needs
  - Satisfy power needs
- Add Fast Initial Response (FIR) feature

Search Algorithm Illustration using the High Side Binomial Case



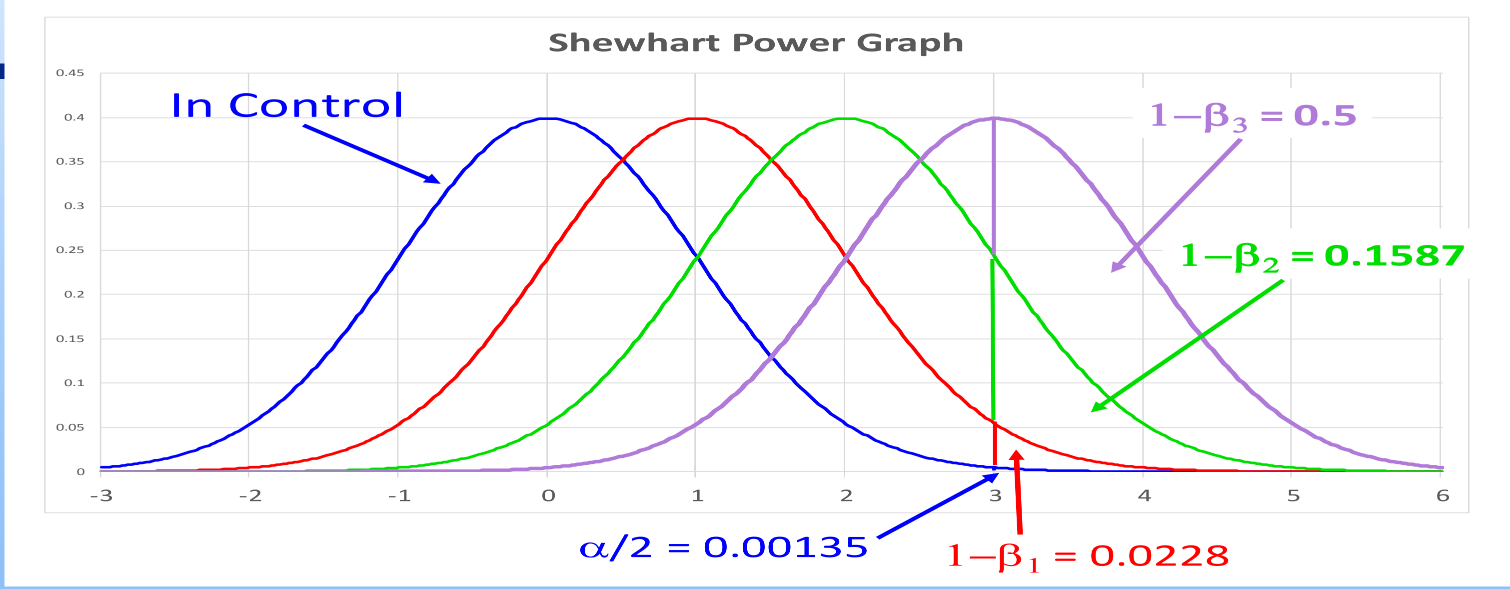
2) Power Considerations for Control Procedures

Power needs lower than in other applications

But…

Power too Low = Poor Control Procedure

Power Considerations: Shewhart



- $ARL_{in\text{-}control} = 1/\alpha = 370$ with $\alpha/2 = 0.00135$
- $ARL_{out\text{-}of\text{-}control} = 1/(1-\beta)$
  - 44 for a 1σ shift because $(1-\beta) = 0.0228$
  - 6.3 for a 2σ shift because $(1-\beta) = 0.1587$
  - 2.0 for a 3σ shift because $(1-\beta) = 0.5$
- Shewhart works well ≥ 2σ shift
- Shewhart works poorly ≤ 1σ shift

Power Considerations – ARL Ratio

Shift	ARL Ratio
1σ	8.4
2σ	59
3σ	185

- ARL Ratio = $ARL_{in\text{-}control}/ARL_{out\text{-}of\text{-}control}$ is used to generalize power to CUSUMs
  - ARL Ratio of 59 works well
  - ARL Ratio of 8.4 is inadequate

- Wide difference between inadequate and good ARL Ratios
- Use CUSUM Experience to tighten the difference

Power Considerations – CUSUM

- Have never seen an effective control scheme with ARL Ratio < 10
- CUSUM with $h = 4$, $k = 0.5$, $S_0=2$ is widely used to detect a 1σ shift

Properties

$ARL_{in\text{-}control}$		$ARL_{out\text{-}of\text{-}control}(\sigma \text{ Shift})$	
0-State	FIR	0-State	FIR
168	149	8.38	5.29

CUSUM Power: Conclusions

- ARL Ratio = $ARL_{in\text{-}control}/ARL_{out\text{-}of\text{-}control}$
- ARL Ratio = 168/8.38 = 20 using 0-State ARLs
- ARL Ratio = 149/8.38 = 18 using FIR_{ic} and 0-$State_{OOC}$ ARLs
- Implications:
  - ARL Ratio < 10 gives poor performance
  - ARL Ratio approaching or >20 usually works well
- Data rich situations can achieve higher ratios

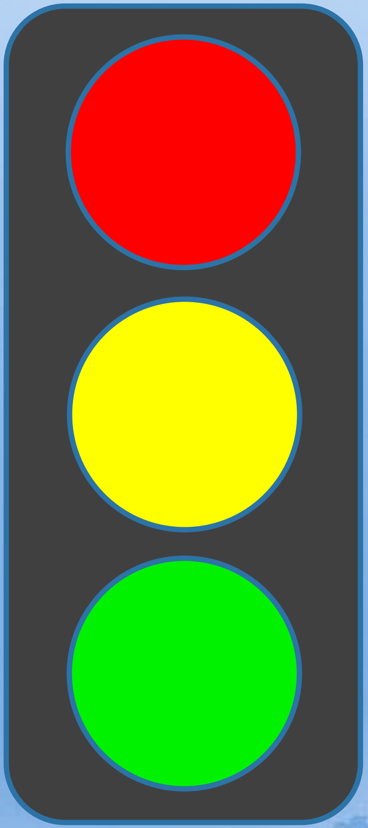
Stoplight Model for Control Scheme Power

$$\text{ARL Ratio} = \text{ARL}_{\text{in-control}} / \text{ARL}_{\text{out-of-control}}$$

$\text{ARL Ratio} < 10$

$10 \leq \text{ARL Ratio} \leq 20$

$20 < \text{ARL Ratio}$



3) Sampling Environments

- Sparse Sampling Environment
  - Short Production Runs
- Standard Sampling Environments
  - ARL within a factor of 3 of a 3σ Shewhart Chart
  - 370/3 ≤ ARL ≤ 3 x 370 = 1110
  - CUSUM design ARL = 500 (often)
- Data Rich Sampling Environment
  - CUSUM design ARL = 5,000

4) The Three Regimes for Counts

See supplemental slides at the end.

- High Count Regime
  - Mean >36
- Intermediate Count Regime
  - 36≥ Mean ≥ 0.04

Discussed next

- Low Count Regime
  - Mean < 0.04

5) Characterizing Control in the Low Count Regime

- High Side Control
  - CUSUM can be effective
  - Can only detect big (percentage) shifts*
  - Usually can not detect doubling*
  - Most Stringent CUSUM control may be required
- Low Side Control
  - Is usually infeasible*
    - $ARL_{in\text{-}control}/ARL_{out\text{-}of\text{-}control}$ Ratio is low
    - In-control ARLs are too large

*Unless you are in a Big Data Environment

Low-Count Control Schemes

- CUSUM – Stringent Control Cases
  - Take Action on Every Count
    - $h < 1.0, \ k = 0.0$
  - 2 Counts in m Observations Signal
    - FIR: A Count in first m Observations Signals
    - CUSUM with $h = 1.0, \ k = 1/m$ and FIR $S_0 = (m-1)/m$
      - This special case uses a larger head start than h/2
- CUSUM: General Case

Low Count- Technical Issues

- For Low Counts Bernoulli, Binomial and Poisson Distributions Converge
- ARLs are close for all distributions
- Bernoulli is Simplest: May consider using Reynolds and Stoumbos (1999)
  - Problems
    - Unstable Formulas
    - Sego L. H. (2006) *Ph.D.Thesis* -- Discussed some of the problems
  - Better Procedure
    - First Permutate Transition Probability Matrix
    - LU Decomposition – Matrix Inversion (Pollard Et Al., 2017)

Take Action on Every Count: Properties

- Most Stringent Control Procedure
- CUSUM Special Case: $h < 1, k=0$
- An Optimum Low Count Procedure
  - Aircraft Accidents and Incidents
  - Benchmarked Industrial Safety System
    - All Accidents are Preventable
- Used if the count is serious enough
  - Serious Events
    - Deepwater Horizon
    - Japan Tsunami
    - Chernobyl
    - Three-Mile Island
- Can be used to judge Quality Awareness
  - How serious must the event be before action is taken

Low Count Regime: m values and the Signal (and Act) on Every Count Procedure

Expected Number of Defects	Bernoulli ARLs*	Optimum Bernoulli m Values (m = 1/k)**						
	Act on Every Count	High Side Shift Detection			Low Side Shift Detection			
		OoM	OoM/2	Doubling	2 OoM	OoM	OoM/2	Halving
0.1	10	3	4	7	**45**	25	20	14
0.05	20	5	8	14	92	51	40	28
0.02	50	12	20	35	**231**	128	100	69
0.01	100	25	40	69	464	255	201	139
0.005	200	51	80	139	929	511	402	277
0.002	500	128	201	347	**2324**	1279	1006	693
0.001	1000	255	402	693	**4650**	2558	2012	1386
0.0005	2000	511	805	1386	9302	5116	4023	2773
0.0002	5000	1279	2012	3466	23257	12792	10059	6931
0.0001	10000	2558	4023	6931	46516	25584	20118	13863
0.00005	20000	5116	8047	13863	**93032**	51168	40236	27726
0.00002	50000	12792	20118	34657	232583	127921	100590	69315
0.00001	100000	25584	40236	69315	**465167**	255842	201180	138629

Note: OoM stands for Order of Magnitude.

* Only the Bernoulli High Side ARL values are shown. All Poisson ARL values are 0.5 greater.

** Only Bernoulli m values are shown as most Poisson values are identical. For the shaded numbers the Poisson m value is one higher. For the bold and darkly shaded numbers, the Poisson value is two higher.

2 in m Control Procedure *

- Allows an occasional count. Signals if 2 counts occur within m samples
- FIR: Signal for 1 count in the first m samples
  - Equivalent to CUSUM with $k = 1/m$; $h=1$
  - FIR head start used is $(m-1)/m$
  - Only case where head start is $> h/2$
- Detects Order of Magnitude Shift
  - $ARL_{in\text{-}control}/ARL_{out\text{-}of\text{-}control}$ Ratio > 20
    - ARLs may be infeasible at Very Low Count Levels
- Poor Performance to detect OOM/2 Shift or to Detect a Doubling
  - $ARL_{in\text{-}control}/ARL_{out\text{-}of\text{-}control}$ Ratio is too low

*Reference: Lucas(1989)

Two Counts in m Observations Signals

Expected # of Defects (np or c)	High Side Control for Detecting a 1 Order of Magnitude Shift							
	In-control ARL		Out-of-control ARL		**FIR** In-control ARL		**FIR** Out-of-control ARL	
	2 in m		**2 in m**		**2 in m**		**2 in m**	
	Bernoulli	Poisson	Bernoulli	Poisson	Bernoulli	Poisson	Bernoulli	Poisson
0.1	62.6	55.6	2.0	2.6	52.6	47.4	1.0	1.7
0.05	127.8	119.7	4.1	4.7	107.8	101.7	2.1	2.8
0.02	300.9	276.5	10.5	10.9	250.9	228.3	5.5	6.0
0.01	566.6	544.8	20.9	21.3	466.6	446.5	10.9	11.4
0.005	1102.2	1096.8	41.7	42.2	902.2	898.5	21.7	22.3
0.002	2727.1	2721.9	104.2	104.7	2227.1	2223.6	54.2	54.8
0.001	5456.2	5435.7	208.4	208.9	4456.2	4437.4	108.4	109.0
0.0005	10883.6	10863.2	416.8	417.3	8883.6	8864.9	216.8	217.3
0.0002	27166.3	27161.1	1042.0	1042.5	22166.3	22162.8	542.0	542.6
0.0001	54319.2	54314.1	2083.9	2084.5	44319.2	44315.8	1083.9	1084.6
0.00005	108625.2	108604.9	4167.9	4168.4	88625.2	88606.6	2167.9	2168.4
0.00002	271512.7	271507.6	10419.6	10420.2	221512.7	221509.3	5419.6	5420.2
0.00001	543012.2	543007.1	20839.2	20839.8	443012.2	443008.8	10839.2	10839.9

Two Counts in m Observations Signals

Expected # of Defects (np or c)	High Side Control for Detecting			
	In-control ARL		Out-of-control ARL	
	2 in m		**2 in m**	
	Bernoulli	Poisson	Bernoulli	Poisson
0.1	62.6	55.6	2.0	2.6
0.05	127.8	119.7	4.1	4.7
0.02	300.9	276.5	10.5	10.9
0.01	566.6	544.8	20.9	21.3

Two Counts in m Observations Signals

Probability of Defect (np)	High Side Control for Detecting a 1/2 Order of Magnitude Shift							
	In-control ARL		Out-of-control ARL		**FIR** In-control ARL		**FIR** Out-of-control ARL	
	2 in m		**2 in m**		**2 in m**		**2 in m**	
	Bernoulli	Poisson	Bernoulli	Poisson	Bernoulli	Poisson	Bernoulli	Poisson
0.1	46.9	44.3	4.3	4.9	36.9	35.5	2.3	3.1
0.05	86.3	84.4	8.6	9.2	66.3	65.6	4.6	5.3
0.02	206.9	205.3	21.6	22.2	156.9	156.4	11.6	12.3
0.01	408.4	406.9	43.1	43.7	308.4	308.0	23.1	23.8
0.005	811.6	810.3	86.3	86.9	611.6	611.3	46.3	46.9
0.002	2015.4	2014.0	215.5	216.1	1515.4	1515.1	115.5	116.1
0.001	4025.9	4024.5	430.9	431.5	3025.9	3025.5	230.9	231.6
0.0005	8040.7	8039.4	861.7	862.3	6040.7	6040.4	461.7	462.4
0.0002	20097.5	20096.2	2154.4	2155.0	15097.5	15097.2	1154.4	1155.0
0.0001	40196.3	40188.9	4308.9	4309.3	30196.3	30189.9	2308.9	2309.4
0.00005	80381.6	80380.3	8617.6	8618.2	60381.6	60381.3	4617.6	4618.3
0.00002	200943.7	200942.4	21544.0	21544.6	150943.7	150943.4	11544.0	11544.6
0.00001	401882.5	401881.2	43088.0	43088.6	301882.5	301882.2	23088.0	23088.6

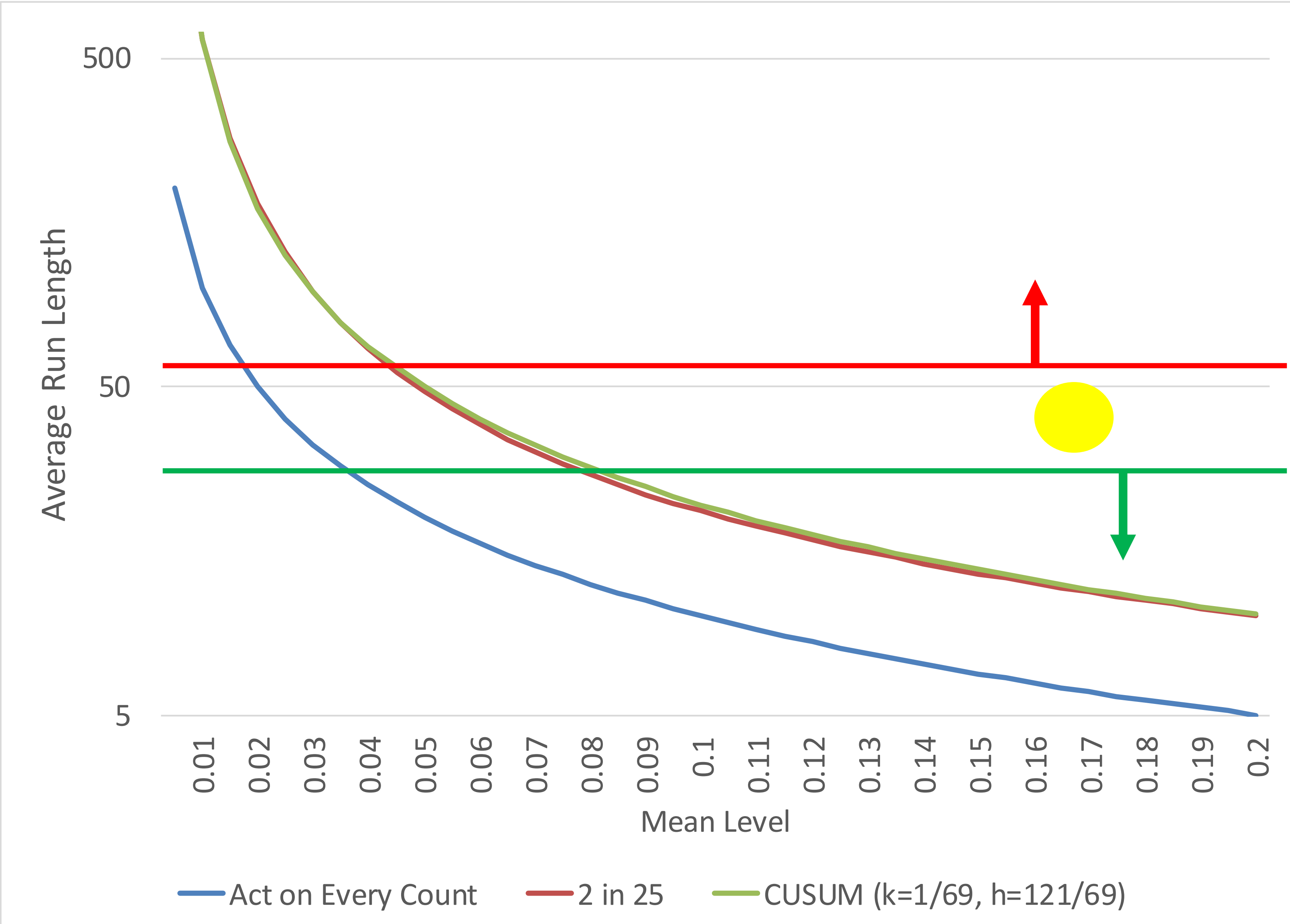
Two Counts in m Observations Signals

Probability of Defect (np)	High Side Control for Detecting a Doubling							
	In-control ARL		Out-of-control ARL		**FIR** In-control ARL		**FIR** Out-of-control ARL	
	2 in m		**2 in m**		**2 in m**		**2 in m**	
	Bernoulli	Poisson	Bernoulli	Poisson	Bernoulli	Poisson	Bernoulli	Poisson
0.1	31.3	31.4	11.8	12.4	21.3	22.0	6.8	7.6
0.05	61.1	61.2	23.4	24.0	41.1	41.8	13.4	14.2
0.02	150.6	150.8	58.3	58.9	100.6	101.3	33.3	34.1
0.01	302.0	302.1	116.9	117.5	202.0	202.7	66.9	67.7
0.005	600.6	600.8	233.3	233.9	400.6	401.3	133.3	134.1
0.002	1500.5	1500.6	583.3	583.9	1000.5	1001.1	333.3	334.0
0.001	3001.6	3001.8	1166.9	1167.5	2001.6	2002.3	666.9	667.6
0.0005	6001.9	6002.1	2333.6	2334.2	4001.9	4002.6	1333.6	1334.4
0.0002	15000.8	15001.0	5833.4	5833.9	10000.8	10001.5	3333.4	3334.1
0.0001	30002.3	30002.4	11667.0	11667.6	20002.3	20002.9	6667.0	6667.8
0.00005	60001.2	60001.4	23333.4	23334.0	40001.2	40001.9	13333.4	13334.2
0.00002	150002.0	150002.2	58333.6	58334.2	100002.0	100002.7	33333.6	33334.4
0.00001	300000.7	300000.9	116666.7	116667.3	200000.7	200001.4	66666.7	66667.4

Standard Sampling Environment: Control action Possibilities when $\mu = 0.01$

- Act on every count
  - Maximizes probability of taking action
- Take Action if 2 counts are too close together
  - 2 in m procedure (CUSUM with h=1, m=1/k=25 or 26)
    - Has adequate power to detect OOM shift
  - CUSUM to detect Doubling (h=121/69, k=1/69)
    - Design for the shift you expect (even if has gives lower power)
- Other Control Procedures have ARL>>500

Control Action Comparison for $\mu_{IC} = 0.01$



Control Action Comparison for $\mu_{IC} = 0.01$

Data plotted on the previous slide

Mean Level	Act on Every Count	2 in 25	CUSUM (k=1/69, h=121/69)
0.01	100.0	566.6	566.1
0.02	50.0	180.1	174.8
0.03	33.3	97.6	96.9
0.04	25.0	65.0	66.1
0.05	20.0	48.2	49.9
0.06	16.7	38.2	39.9
0.07	14.3	31.6	33.2
0.08	12.5	27.0	28.4
0.09	11.1	23.5	24.7
0.1	10.0	20.9	21.9
0.11	9.1	18.8	19.6
0.12	8.3	17.1	17.7
0.13	7.7	15.7	16.2
0.14	7.1	14.5	14.9
0.15	6.7	13.5	13.8
0.16	6.3	12.6	12.9
0.17	5.9	11.8	12.1
0.18	5.6	11.2	11.3
0.19	5.3	10.6	10.7
0.2	5.0	10.0	10.1

CUSUM with k = 1/m, h = 2-1/m

- Transition Matrix of Poisson, Binomial and Bernoulli have the same form
  - Only use P(0) and P(1)
  - Two counts is an OOC Signal
- Can detect OOM/2 shift
  - ARL is too large for very low count levels
- ARL Ratio (Power) too low to detect a doubling

CUSUM with k = 1/m, h = 2-1/m

	High Side Control for Detecting a 1 Order of Magnitude Shift									
	Zero State				FIR State					
	In-control ARL		Out-of-control ARL		In-control ARL		Out-of-control ARL		Zero State	
Expected #	CUSUM		CUSUM		CUSUM		CUSUM		ARL Ratio = IC/OOC	
of Defects	h = 2-1/m	k=1/m	h = 2-1/m	k=1/m	h = 2-1/m	k=1/m	h = 2-1/m	k=1/m		
(np or c)	Bernoulli	Poisson	Bernoulli	Poisson	Bernoulli	Poisson	Bernoulli	Poisson	Bernoulli	Poisson
0.1	**377.1**	**150.0**	**3.0**	**3.2**	**367.1**	**146.7**	**2.0**	**2.7**	**126**	**48**
0.04	1209.9	659.8	8.3	8.3	1184.9	647.2	5.8	6.5	146	79
0.02	2263.8	1470.2	16.9	16.6	2213.8	1433.9	11.9	12.3	134	89
0.01	**4082.1**	**3275.6**	**33.7**	**33.4**	**3982.1**	**3191.1**	**23.7**	**24.1**	**121**	**98**
0.005	7777.5	7155.2	67.4	67.4	7577.5	6972.9	47.4	48.1	115	106
0.004	9630.3	9036.2	84.2	84.2	9380.3	8803.3	59.2	59.9	114	107
0.002	19162.1	18555.8	168.7	168.7	18662.1	18073.4	118.7	119.4	114	110
0.001	**38488.2**	**36923.1**	**338.1**	**337.7**	**37488.2**	**35959.1**	**238.1**	**238.4**	**114**	**109**
0.0005	76616.9	76356.8	676.0	675.7	74616.9	74356.8	476.0	476.4	113	113
0.0004	95681.6	95421.9	845.0	844.7	93181.6	92921.9	595.0	595.4	113	113
0.0002	191006.5	191006.5	1689.9	1690.0	186006.5	186006.5	1189.9	1190.7	113	113
0.0001	**381916.4**	**381916.4**	**3380.2**	**3380.2**	**371916.4**	**371916.4**	**2380.2**	**2380.9**	**113**	**113**
0.00005	763736.2	763476.9	6760.6	6760.3	743736.2	743476.9	4760.6	4761.0	113	113
0.00004	954386.9	954386.9	8450.5	8450.5	929386.9	929386.9	5950.5	5951.2	113	113
0.00002	1908677.3	1908677.3	16901.2	16901.3	1858677.3	1858677.3	11901.2	11901.9	113	113
0.00001	3817258.1	3817258.1	33802.7	33802.8	3717258.1	3717258.1	23802.7	23803.5	113	113

CUSUM with k = 1/m, h = 2-1/m

	High Side Control for Detecting a 1/2 Order of Magnitude Shift									
	Zero State				FIR State					
	In-control ARL		Out-of-control ARL		In-control ARL		Out-of-control ARL		Zero State	
Expected #	CUSUM		CUSUM		CUSUM		CUSUM		ARL Ratio = IC/OOC	
of Defects	h = 2-1/m	k=1/m	h = 2-1/m	k=1/m	h = 2-1/m	k=1/m	h = 2-1/m	k=1/m		
(np or c)	Bernoulli	Poisson	Bernoulli	Poisson	Bernoulli	Poisson	Bernoulli	Poisson	Bernoulli	Poisson
0.1	**211.2**	**121.0**	**6.9**	**6.9**	**201.2**	**116.2**	**4.9**	**5.7**	**31**	**17**
0.04	479.9	377.7	18.0	17.9	454.9	359.4	13.0	13.7	27	21
0.02	934.7	826.2	36.4	36.3	884.7	783.5	26.4	27.1	26	23
0.01	**1846.5**	**1732.9**	**73.3**	**73.1**	**1746.5**	**1640.6**	**53.3**	**54.0**	**25**	**24**
0.005	3670.8	3549.2	147.0	146.8	3470.8	3357.5	107.0	107.6	25	24
0.004	4583.1	4400.9	183.8	182.9	4333.1	4158.8	133.8	133.7	25	24
0.002	9073.9	8959.5	367.2	367.1	8573.9	8467.5	267.2	267.9	25	24
0.001	**18126.8**	**17853.8**	**734.8**	**734.7**	**17126.8**	**16870.6**	**534.8**	**535.5**	**25**	**24**
0.0005	36163.0	36163.0	1469.2	1469.1	34163.0	34163.0	1069.2	1069.9	25	25
0.0004	45215.9	45215.9	1836.8	1836.7	42715.9	42715.9	1336.8	1337.5	25	25
0.0002	90411.0	90411.0	3674.0	3673.9	85411.0	85411.0	2674.0	2674.7	25	25
0.0001	**180870.9**	**180801.3**	**7349.2**	**7348.3**	**170870.9**	**170801.3**	**5349.2**	**5349.1**	**25**	**25**
0.00005	361651.4	361651.4	14697.9	14697.8	341651.4	341651.4	10697.9	10698.6	25	25
0.00004	452041.6	452041.6	18372.3	18372.2	427041.6	427041.6	13372.3	13373.0	25	25
0.00002	904062.4	904062.4	36745.0	36744.9	854062.4	854062.4	26745.0	26745.7	25	25
0.00001	1808104.0	1808104.0	73490.4	73489.2	1708104.0	1708104.0	53490.4	53490.4	25	25

CUSUM with k = 1/m, h = 2-1/m

Expected # of Defects (np or c)	High Side Control for Detecting a Doubling								Zero State ARL Ratio = IC/OOC	
	Zero State				FIR State					
	In-control ARL		Out-of-control ARL		In-control ARL		Out-of-control ARL			
	CUSUM h = 2-1/m k=1/m		CUSUM h = 2-1/m k=1/m		CUSUM h = 2-1/m k=1/m		CUSUM h = 2-1/m k=1/m			
	Bernoulli	Poisson	Bernoulli	Poisson	Bernoulli	Poisson	Bernoulli	Poisson	Bernoulli	Poisson
0.1	**86.8**	**73.2**	**21.8**	**20.9**	**76.8**	**65.9**	**16.8**	**17.1**	**4**	**3**
0.04	220.6	204.6	56.3	55.2	195.6	182.8	43.8	43.9	4	4
0.02	423.1	408.5	110.8	109.8	373.1	361.6	85.8	86.0	4	4
0.01	**858.8**	**843.0**	**224.2**	**223.1**	**758.8**	**746.2**	**174.2**	**174.3**	**4**	**4**
0.005	1700.4	1683.6	446.7	445.6	1500.4	1487.0	346.7	346.8	4	4
0.004	2136.0	2120.3	560.0	558.9	1886.0	1873.5	435.0	435.2	4	4
0.002	4254.9	4239.3	1118.4	1117.3	3754.9	3742.6	868.4	868.5	4	4
0.001	**8522.2**	**8472.4**	**2239.3**	**2238.2**	**7522.2**	**7479.9**	**1739.3**	**1739.4**	**4**	**4**
0.0005	17042.2	17042.2	4479.0	4477.9	15042.2	15042.2	3479.0	3479.1	4	4
0.0004	21294.8	21294.8	5597.8	5596.7	18794.8	18794.8	4347.8	4347.9	4	4
0.0002	42587.3	42587.3	11195.9	11194.8	37587.3	37587.3	8695.9	8696.1	4	4
0.0001	**85187.0**	**85187.0**	**22394.4**	**22393.3**	**75187.0**	**75187.0**	**17394.4**	**17394.6**	**4**	**4**
0.00005	170357.1	170357.1	44787.1	44786.0	150357.1	150357.1	34787.1	34787.3	4	4
0.00004	212942.1	212942.1	55983.5	55981.9	187942.1	187942.1	43483.5	43483.2	4	4
0.00002	425896.7	425896.7	111969.5	111970.6	375896.7	375896.7	86969.5	86971.2	4	4
0.00001	851776.4	851776.4	223937.3	223927.7	751776.4	751776.4	173937.3	173931.2	4	4

Detecting Doubling

- Intermediate Regime (Mean ≥ 0.04)
  - Possible in Data Rich Sampling Environment (h = 5 or 6 need for adequate power)
- Low Count Regime (Mean < 0.04)
  - ARL Ratio too low*
  - $\text{ARL}_{\text{in-control}}$ too large*
    - *Except for a Big Data Environment

Detecting Doubling Needs Larger h Values Bernoulli CUSUMS with Larger h Values

h	ARL Ratio*	$(ARL_{ooc})*(\mu_{ooc})$	$ARL_{ic}(\mu= 0.01)$	$ARL_{ooc}(\mu= 0.02)$
2	3.9-3.8	4.5-4.5	868.8	225.6
3	6.2-6.0	7.5-7.4	2,300.7	373.4
4	10.2-9.8	10.8-10.7	5,516.8	541.0
5	17.2-16.4	14.3-14.1	12,323.9	716.1
6	29.5-27.7	17.9-17.7	26,409.6	894.5

Table shows that detecting doubling in Standard and Data Rich Sampling Environments is unfeasible for $\mu \leq 0.01$)

The average number of counts needed to signal is $(ARL_{ooc})*(\mu_{ooc})$

We are conservative and recommend a “Rule of 20” to Detect Doubling

* The ARL Ratio (= ARL_{ic}/ARL_{ooc}) is for one sided control

High Side Control for Detecting a 1 Order of Magnitude Shift

Probability of Defect (np)	Bernoulli CUSUM with **h=5 k=1/t**				Bernoulli CUSUM with **h=6 k=1/t**			
	Zero-State ($S_0 = 0$)		FIR ($S_0 = 0.5*hm$)		Zero-State ($S_0 = 0$)		FIR ($S_0 = 0.5*hm$)	
	In-Control ARL	Out-of-Control ARL	In-Control ARL	Out-of-Control ARL	In-Control ARL	Out-of-Control ARL	In-Control ARL	Out-of-Control ARL
0.01	4972156.4	82.6	4966532.8	47.9	55446146.0	99.3	55426289.7	55.9
0.001	32773000.5	814.4	32725392.0	473.9	332371578.4	979.2	332219228.6	556.4
0.0001	312894513.2	8128.0	312426624.7	4729.5	3139016894.6	9772.1	3137535329.7	5557.4

High Side Control for Detecting a 1/2 Order of Magnitude Shift

Probability of Defect (np)	Bernoulli CUSUM with **h=5 k=1/t**				Bernoulli CUSUM with **h=6 k=1/t**			
	Zero-State ($S_0 = 0$)		FIR ($S_0 = 0.5*hm$)		Zero-State ($S_0 = 0$)		FIR ($S_0 = 0.5*hm$)	
	In-Control ARL	Out-of-Control ARL	In-Control ARL	Out-of-Control ARL	In-Control ARL	Out-of-Control ARL	In-Control ARL	Out-of-Control ARL
0.01	264805.8	190.5	262447.7	113.8	1384144.0	230.5	1378534.1	134.0
0.001	2250296.7	1891.2	2228720.4	1137.7	11328923.7	2289.5	11278462.5	1339.6
0.0001	22117730.3	18895.4	21903958.9	11374.1	110898212.6	22874.8	110399060.7	13391.8

CUSUMS with h = 5 & 6

ARLs often in a Big Data Sampling Environment for OOM & OOM/2 Shifts

CUSUMS with h = 5 & 6

Detect Doubling with an ARL ratio > 20 for h = 6

High Side Control for Detecting a Doubling

Probability of Defect (np)	Bernoulli CUSUM with **h=5 k=1/t**				Bernoulli CUSUM with **h=6 k=1/t**			
	Zero-State ($S_0 = 0$)		FIR ($S_0 = 0.5$*hm)		Zero-State ($S_0 = 0$)		FIR ($S_0 = 0.5$*hm)	
	In-Control ARL	Out-of-Control ARL	In-Control ARL	Out-of-Control ARL	In-Control ARL	Out-of-Control ARL	In-Control ARL	Out-of-Control ARL
0.1	1690.1	72.9	1569.8	47.3	3825.8	90.3	3602.0	55.6
0.04	3766.5	186.0	3496.9	123.9	8469.9	232.4	7991.6	148.9
0.02	5973.5	351.8	5481.1	232.8	12621.9	438.2	11781.4	280.0
0.01	12323.9	716.1	11342.3	477.6	26409.6	894.5	24737.9	577.5
0.005	23349.5	1413.1	21427.3	940.8	49312.8	1763.6	46068.4	1137.8
0.004	29679.9	1777.3	27268.8	1185.5	63037.4	2219.9	58962.8	1435.3
0.002	58105.5	3535.6	53323.4	2356.7	122702.4	4414.6	114649.3	2853.5
0.001	116560.3	7083.5	106999.6	4725.3	246481.4	8847.2	230385.4	5724.6
0.0005	232677.0	14163.6	213569.4	9448.4	491841.5	17690.8	459702.1	11449.3
0.0004	290340.5	17695.9	266480.0	11805.0	613427.6	22101.9	573287.4	14303.1
0.0002	580239.6	35388.6	532532.4	23607.7	1225740.7	44200.2	1145512.6	28606.3
0.0001	1160826.3	70789.4	1065431.0	47229.5	2452551.5	88418.4	2292105.6	57230.3
0.00005	2320423.4	141559.9	2129672.0	94444.8	4901805.4	176811.8	4581007.9	114443.4
0.00004	2900222.0	176945.2	2661792.6	118052.5	6126432.9	221008.5	5725459.5	143050.0
0.00002	5800791.0	353902.6	5323935.8	236116.8	12253935.1	442034.9	11451998.5	286117.5
0.00001	11600353.3	707786.3	10646682.3	472219.4	24504574.9	884044.8	22900795.9	572218.0

6) More About Detecting Doubling

- Two Armed Clinical Trial
  - Rule of 50
- Control Schemes
  - Rule of 20
- Intermediate Count Regime
  - $\mu = 0.1$ & 0.04
  - Two sided control Feasible
    - Detecting halving – ARL Ratio too low
    - Detect OOM/2 Improvement – ARL Ratio Good
- Standard Sampling Environment
  - How large must μ be to Detect Doubling

Clinical Trials: Rule of 50

Statistical Power (%) to Detect a Doubling of Adverse Event Rates in Clinical Studies of Drugs, by Sample Size

Sample Size	From 5% to 10%, %	From 1% to 2%, %	From 0.1% to 0.2%, %
1 000	82	17	5
5 000	>99	80	7
10 000	>99	>98	17
50 000	>99	>99	79

Am J Public Health. 2008 August; 98(8): 1366–1371.

All cases consider a clinical trial with two equally sized groups.
The cases with a power around 80% show counts of 25 and 50.
This has been described as a rule of 50

Rule of 20: Additional Reasons

- When is an Observed Doubling Significant
- Pearson and Hartley Table 36
- Test for the Significance of the Difference Between two Poisson Variables: one-tailed test

Higher Count	Lower Count	Significance level
14	7	0.10
20	10	0.05
28	14	0.025
36	18	0.01
44	22	0.005

Detecting Doubling for μ = 0.1 & 0.04

(for 2-sided control, low-side is an OOM/2 shift)

Goal Mean	Spec Arl	High Side Control h	ARL (μ)	ARL (2μ)	ARL Ratio	Stop Light	OOC Ct	Low Side Control h	ARL (μ)	ARL (μ/5)	ARL Ratio	Stop Light
0.04	500	2.76	500.59	88.41	5.7	Red	7.1	1.55	516.1	104.9	4.9	Red
0.04	1,000	3.5	1,004.84	120.7	8.3	Red	9.7	1.93	1,034.10	135.5	7.6	Red
0.04	2,000	4.3	2,030.79	157.7	12.9	Yellow	12.6	2.33	2,068.13	168.2	12.3	Yellow
0.04	3,000	4.78	3,034.64	180.5	16.8	Yellow	14.4	2.57	3097.88	188	16.5	Yellow
0.04	5,000	5.4	5,037.03	210.4	23.9	Green	16.8	2.87	5,092.30	212.8	23.9	Green
0.04	10,000	6.28	10,141.10	253.4	40	Green	20.3	3.29	10,109.50	247.7	40.8	Green
0.1	500	3.94	505.57	53.17	9.5	Red	10.6	2.025	515	58	8.9	Red
0.1	1,000	4.86	1,001.06	67.99	14.7	Yellow	13.6	2.425	1021.4	71.1	14.4	Yellow
0.1	2,000	5.86	2,026.49	84.49	24	Green	16.9	2.875	2156.72	86	25.1	Green
0.1	3,000	6.44	3,007.55	94.08	32	Green	18.8	3.125	3245.99	94.29	34.4	Green
0.1	5,000	7.22	5,054.66	107	47.2	Green	21.4	3.425	5281.2	104.3	50.6	Green
0.1	10,000	8.26	10,020.30	124.3	80.6	Green	24.9	3.825	10062.5	117.6	85.6	Green

High Side Control	Low Side Control
Mean = 0.04 k = 0.0577 (used 0.06)	Mean = 0.04 k = 0.0199 (used 0.02)
Mean = 0.10 k = 0.14427 (used 0.14)	Mean = 0.10 k = 0.0497 (used 0.05)

Detect Doubling – Standard Sampling Environment

Goal Mean	Spec Arl	High Side Control h	ARL (μ)	ARL (2μ)	ARL Ratio	Stop Light	OOC Ct	Low Side Control h	ARL (μ)	ARL (μ/5)	ARL Ratio	Stop Light
0.5	500	5.42	507.6	18.25	27.8	Green	18.3	2.66	596.29	18.14	32.9	Green
1	500	5.94	505.40	10.93	46.2	Green	21.9	3.25	801.10	10.89	73.6	Green
0.5	1,000	6.38	1,000.13	21.58	46.3	Green	21.6	3.3	1,336.70	21.45	62.3	Green
1	1,000	6.94	1,019.13	12.69	80.3	Green	25.3	3.75	1788.44	12.56	142.4	Green

High Side Control	Low Side Control
Mean = 0.5 k = 0.7213 (used 0.72)	Mean = 0.5 k = 0.2485 (used 0.25)
Mean = 1 k = 1.4427 (used 1.44)	Mean = 1 k = 0.4971 (used 0.5)

By the time the goal mean is 0.5 and ARL>1000 (2-sided ARL > 500):

- Doubling can be detected on the high side
- An OoM/2 shift can be detected on the low side

Detecting Improvement

- Usually not possible in the Low Count Regime
- Detecting 2 (OOM) shift shown on next slide
- ARLs usually too large in the Low Count Regime unless you are in a Very Data Rich Sampling Environment

- Detecting improvement is possible in the Intermediate Count Regime ($\mu \geq 0.04$)

Detect 2 (OOM) Shifts Using m Zeros in a Row Rule

Expected # of Defects (np or c)	Low Side Control for Detecting a 2 Orders of Magnitude Shift			
	In-control ARL		Out-of-control ARL	
	M in a Row h=1 k=1/m		**M in a Row** h=1 k=1/m	
	Bernoulli	Poisson	Bernoulli	Poisson
0.5	510.0	226.2	8.2	9.2
0.1	1135.7	1144.9	46.1	48.1
0.04	2708.7	2615.2	117.7	118.8
0.02	5268.0	5284.3	236.4	238.5
0.01	10499.1	10410.4	475.0	476.0
0.005	20856.6	20768.8	950.9	952.0
0.004	26088.4	26000.5	1189.5	1190.5
0.002	51931.5	51949.1	2378.9	2381.0
0.001	103828.6	103846.3	4759.8	4761.9
0.0005	207622.9	207535.8	9521.7	9522.8
0.0004	259520.1	259433.0	11902.7	11903.7
0.0002	518901.2	518814.2	23806.4	23807.4

7) Conclusions

- Control
  - Less power required than for other applications
    - However, Too Low Power = Poor Control Procedure
- Characterizing the Low Count Regime
  - Sampling Environment (Standard, Data Rich, Big Data)
  - ARL Ratio (Stoplight Model)
  - Shift Size (OOM, OOM/2, Doubling)
- Low Count Regime
  - Act on Every count when $\mu < 0.001$ (or $\mu < 0.0001$)
  - 2-in-m Procedure for OOM shifts
  - $h = 2 - 1/m$ CUSUM for OOM/2 shifts
  - Detecting Doubling or Improvement is not feasible in the low count regime (except in Big Data Environments)
  - Doubling can be detected in the Intermediate Count Regime
  - Doubling Detected with $\mu \geq 0.5$ in Standard Sampling Environment

Thank You !

Supplemental Slides

References

- Gan, F.F. (1993) "An optimum design of CUSUM control charts for binomial counts" *Journal of Applied Statistics* **11**, 16–31.
- Lucas, J.M. (1976), "The Design and Use of V-Mask Control Schemes", Journal of Quality Technology, Volume 8, pages 1-12 .
- Lucas, J. M. and R. B. Crosier (1982a), "Fast Initial Response for CUSUM Quality Control Schemes: Give Your CUSUM a Head Start", Technometrics, Volume 24, pages 199-205.
- Lucas, J. M. (1985a), "Counted Data CUSUMs", Technometrics, Volume 27, pages 129 -144.
- Lucas, J.M. (1989) "Control schemes for low count levels" *Journal of Quality Technology*, **21**, 199–201.
- Lucas, J.M. and Saccucci, M.S. (1990), "Exponentially weighted moving average control schemes: properties and enhancements" *Technometrics*,**32**, 1–12.
- Moustakides GV. (1986) "Optimal stopping times for detecting changes in distributions." *Annals of Statistics* **14**:1379–1387.
- Lucas, Davis & Saniga (2006) "Detecting Improvement with Shewhart Attribute Control Charts When the Lower Control Limit is Zero," *IIE Transactions*, Vol. 38, No. 8, pp. 659-669
- Reynolds Jr MR, Stoumbos ZG.(1999) "A CUSUM chart for monitoring a proportion when inspecting continuously" *Journal of Quality Technology* 1999; **31**:87–108.
- A Review and Perspective on Surveillance of Bernoulli ProcessesbQual. Reliab. Engng. Int. 2011, 27 735–752
- Landon H. Sego1, William H. Woodall and Marion R. Reynolds Jr, (2008)"A comparison of surveillance methods for small incidence rates" *STATISTICS IN MEDICINE* **27**:1225–1247
- Pearson,E. S., and Hartley H. O. (1962) "Biometrika Tables for Statisticians" Cambridge University Press
- Sego L. H. (2006) *Ph.D. Thesis*, Virginia Polytechnic Institute and State University, Blacksburg, VA, http://scholar.lib.vt.edu/theses/available/etd-04122006-182352/unrestricted/Landon Sego dissertation.pdf
- Woodall WH, Mahmoud MA. (2005), "The inertial properties of quality control charts" *Technometrics* 2005; **47**:425–443.
- Woodall WH. (1997) "Control charts based on attribute data: bibliography and review." *Journal of Quality Technology*1997; **29**:172–182.

2 in m versus Act on every count

- Note that (Exp. # of Defects)*(ARLin-control) changes smoothly from 6.26 to 5.43
- Generalization: When the process mean is $5.5/ARL_{in\text{-}control}$, then you can use the 2 in m procedure to detect an OOM shift (to $55/ARL_{in\text{-}control}$).
- If smaller shifts need to be detected, then the "take control action on every count" procedure should be considered

CUSUMS with h = 5 & 6

ARLs too large for OOM & OOM/2 Shifts

Detects Doubling for largest means shown

Enlarged on next 2 slides

High Side Control for Detecting a 1 Order of Magnitude Shift

Probability of Defect (np)	Bernoulli CUSUM with **h=5 k=1/t**				Bernoulli CUSUM with **h=6 k=1/t**			
	Zero-State ($S_0 = 0$)		FIR ($S_0 = 0.5$*hm)		Zero-State ($S_0 = 0$)		FIR ($S_0 = 0.5$*hm)	
	In-Control ARL	Out-of-Control ARL	In-Control ARL	Out-of-Control ARL	In-Control ARL	Out-of-Control ARL	In-Control ARL	Out-of-Control ARL
0.01	4972156.4	82.6	4966532.8	47.9	55446146.0	99.3	55426289.7	55.9
0.001	32773000.5	814.4	32725392.0	473.9	332371578.4	979.2	332219228.6	556.4
0.0001	312894513.2	8128.0	312426624.7	4729.5	3139016894.6	9772.1	3137535329.7	5557.4

High Side Control for Detecting a 1/2 Order of Magnitude Shift

Probability of Defect (np)	Bernoulli CUSUM with **h=5 k=1/t**				Bernoulli CUSUM with **h=6 k=1/t**			
	Zero-State ($S_0 = 0$)		FIR ($S_0 = 0.5$*hm)		Zero-State ($S_0 = 0$)		FIR ($S_0 = 0.5$*hm)	
	In-Control ARL	Out-of-Control ARL	In-Control ARL	Out-of-Control ARL	In-Control ARL	Out-of-Control ARL	In-Control ARL	Out-of-Control ARL
0.01	264805.8	190.5	262447.7	113.8	1384144.0	230.5	1378534.1	134.0
0.001	2250296.7	1891.2	2228720.4	1137.7	11328923.7	2289.5	11278462.5	1339.6
0.0001	22117730.3	18895.4	21903958.9	11374.1	110898212.6	22874.8	110399060.7	13391.8

High Side Control for Detecting a Doubling

Probability of Defect (np)	Bernoulli CUSUM with **h=5 k=1/t**				Bernoulli CUSUM with **h=6 k=1/t**			
	Zero-State ($S_0 = 0$)		FIR ($S_0 = 0.5$*hm)		Zero-State ($S_0 = 0$)		FIR ($S_0 = 0.5$*hm)	
	In-Control ARL	Out-of-Control ARL	In-Control ARL	Out-of-Control ARL	In-Control ARL	Out-of-Control ARL	In-Control ARL	Out-of-Control ARL
0.1	1690.1	72.9	1569.8	47.3	3825.8	90.3	3602.0	55.6
0.04	3766.5	186.0	3496.9	123.9	8469.9	232.4	7991.6	148.9
0.02	5973.5	351.8	5481.1	232.8	12621.9	438.2	11781.4	280.0
0.01	12323.9	716.1	11342.3	477.6	26409.6	894.5	24737.9	577.5
0.005	23349.5	1413.1	21427.3	940.8	49312.8	1763.6	46068.4	1137.8
0.004	29679.9	1777.3	27268.8	1185.5	63037.4	2219.9	58962.8	1435.3
0.002	58105.5	3535.6	53323.4	2356.7	122702.4	4414.6	114649.3	2853.5
0.001	116560.3	7083.5	106999.6	4725.3	246481.4	8847.2	230385.4	5724.6
0.0005	232677.0	14163.6	213569.4	9448.4	491841.5	17690.8	459702.1	11449.3
0.0004	290340.5	17695.9	266480.0	11805.0	613427.6	22101.9	573287.4	14303.1
0.0002	580239.6	35388.6	532532.4	23607.7	1225740.7	44200.2	1145512.6	28606.3
0.0001	1160826.3	70789.4	1065431.0	47229.5	2452551.5	88418.4	2292105.6	57230.3
0.00005	2320423.4	141559.9	2129672.0	94444.8	4901805.4	176811.8	4581007.9	114443.4
0.00004	2900222.0	176945.2	2661792.6	118052.5	6126432.9	221008.5	5725459.5	143050.0
0.00002	5800791.0	353902.6	5323935.8	236116.8	12253935.1	442034.9	11451998.5	286117.5
0.00001	11600353.3	707786.3	10646682.3	472219.4	24504574.9	884044.8	22900795.9	572218.0

Short Runs Problem

- Sparse Data because there are only a few samples taken before the run is completed
- Use CUSUM with a FIR Feature
  - Shorter in-control ARL (Use smaller h)
  - Detect Larger Shifts (Use larger k)
- Alternative approaches use EWMA with a FIR Feature

Short Runs: Some CUSUM Options

Average Run Lengths (FIR ARLs) at shift:

Shift	k	h	0.0	1.0	1.5	2.0	ARL Ratio	Stop Light
1.0	0.5	4.0	168 (149)	8.4 (5.3)	4.8 (2.9)	3.3 (2.0)	20.0	Green
1.0	0.5	3.5	98 (84)	7.4 (4.7)	4.3 (2.6)	3.0 (1.8)	13.2	Yellow
1.0	0.5	3.0	59 (49)	6.4 (4.2)	3.8 (2.4)	2.7 (1.7)	9.2	Red
1.5	0.75	3.0	221 (205)	9.7 (6.8)	4.7 (3.0)	3.1 (2.0)	46.7	Green
1.5	0.75	2.25	70 (62)	7.1 (5.1)	3.7 (2.5)	2.5 (1.7)	18.9	Yellow
1.5	0.75	1.5	21 (18)	4.8 (3.7)	2.7 (2.0)	1.9 (1.4)	7.8	Red

Short Runs: Diagnostic ANOVA

- The Analysis of Variance should include the following terms:
  - Run to Run (σ_{Run})
  - Shifts Within Runs (σ_{Shift})
    - Use indicator Variables
  - Sample to Sample (σ_{Sample})
  - Replicates (σ_{reps})
- Projects to reduce the Run to Run component (σ_{Run}) are often undertaken

Signal at or beyond parameter value?

- Shewhart Special Case:
  - k = Shewhart Limit; h = 0
  - Must exceed Shewhart Limit to signal (our convention)
- CUSUM General Case
  - k > 0; h > 0
  - For Integer k,
    - Markov Chain: (h) x (k) states
    - Indexed 0, 1, … hk-1
  - Signal at h because the state indexed hk-1 is the highest in-control state

High-Count Regime

- Count level greater than 8 to 16*; we now say ≥ 36
- Two sided control is almost always used.
- The control scheme can be symmetrical (and look the same for both the high and low sides when a V-mask is used).
- Only integer arithmetic needs to be used (because there can be almost no improvement from using non-integer parameter values).
- Easy to detect doubling or level cut in half.
  - Even with a Shewhart Chart because $C + 3(C)^{1/2} < 2C$ when $C>9$ and $C - 3(C)^{1/2} > C/2$ for $C>36$
  - Doubling is (usually) larger than a 1σ shift (so a Shewhart Chart can usually detect a doubling).
- CUSUM for Counts ARL Curcves are very similar to CUSUM for variables curves
- For average >100 we can use square Root transformation and Variable ARL curves even though our program handles, without problems larger averages
- In Standard ARL situations can very effectively detect 1σ shift.
  - Standard CUSUM can usually effectively detect a shift $>\sigma/2$.
- In enhanced data situations smaller shifts can be detected.

* DuPont PQM Manual Recommendation

Intermediate Count Regime

- Count average between 2 to 4 and 8 to 16* We now say .04 to 36
- Two sided CUSUM control can be used effectively.
- Shewhart Lower Limit is often negative.

  Count the zeros can be used to detect low-side shifts
- Appropriate CUSUM is often un-symmetric when a V-mask is used.
- Low-side (in-control) ARLs often chosen larger than high-side ARLs.
- In Standard ARL situations very can effectively detect 1σ shift.
  - Standard CUSUM can effectively detect a shift >σ/2.
    - 1σ shift can be greater than doubling.
- Non integer parameter values can give improved properties.
- In enhanced data situations smaller shifts can be detected.

* DuPont PQM Manual Recommendation

Design of CUSUM for Counts

S_i = CUSUM at stage i
y_i = the observed count
k = reference value
h = decision interval

For the high side

$S_i = \text{Max}\{0, S_{i-1} + y_i - k_H\}$ signal if $S_i \geq h_H$

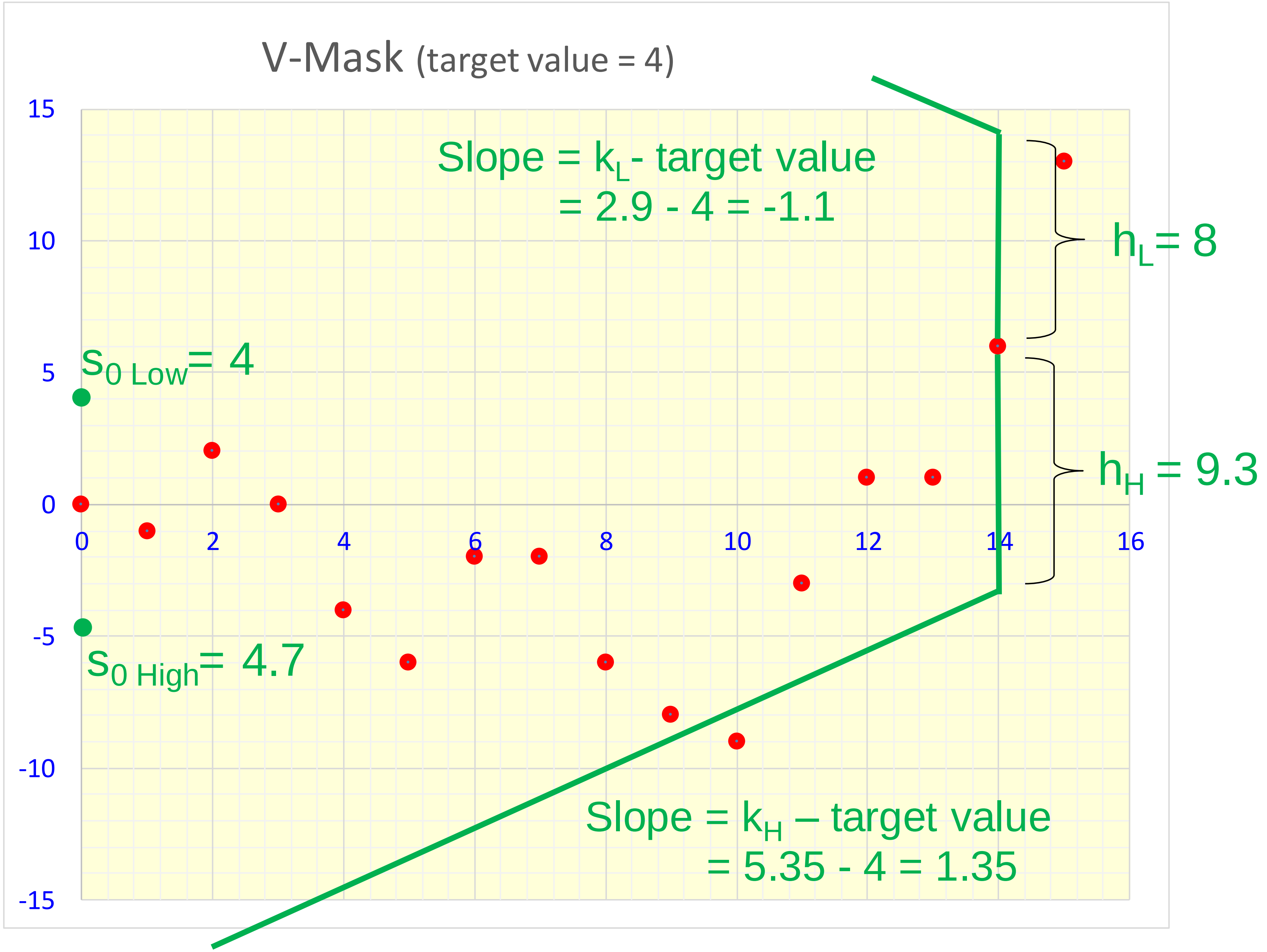
For the low side

$S_i = \text{Max}\{0, S_{i-1} + k_L - y_i\}$ signal if $S_i \geq h_L$

Our objective: An automatic design algorithm to find k_H, h_H, k_L, and h_L values that meet the user's in-control and out-of-control ARL needs.

Automatic Design of CUSUMs for Counts

- Multi-stage project in progress…
  - Design for Binomial, Poisson and Bernoulli
  - Algorithm illustrated on next slide – selecting ***k*** and ***h*** to satisfy ARL needs. If **R** represents the transition probability matrix for the non-signaling S_i values, then ARL vector is: $\mathbf{ARL} = (\mathbf{I} - \mathbf{R})^{-1}\,(\mathbf{1})$
    - 1st element is the ARL zero state
    - "Middle" element is the FIR Head Start ARL Value



V-Mask (target value = 4)
Slope = kL- target value
= 2.9 - 4 = -1.1
hL= 8
hH = 9.3
s0 Low= 4
s0 High= 4.7
Slope = kH – target value
= 5.35 - 4 = 1.35
15
10
5
0
-5
-10
-15
0
2
4
6
8
10
12
14
16