SAND2023-05824

SANDIA REPORT
SAND2023-05824
Printed July 2023

Statistical Framework for Planning a Shelf Life Program with Binary Performance Data

Stephen V. Crowder

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico
87185 and Livermore,
California 94550

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology & Engineering Solutions of Sandia, LLC.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone:	(865) 576-8401
Facsimile:	(865) 576-5728
E-Mail:	reports@osti.gov
Online ordering:	http://www.osti.gov/scitech

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone:	(800) 553-6847
Facsimile:	(703) 605-6900
E-Mail:	orders@ntis.gov
Online order:	https://classic.ntis.gov/help/order-methods/

ABSTRACT

This document outlines a statistical framework for establishing a shelf-life program for components whose performance is measured by a binary response, usually 'pass' or 'fail.' The approach applies to both single measurement devices and repeated measurement devices. The high-level objective of these plans is to quickly detect any sizeable increase in fraction defective as the product ages. The statistical approach is to choose a sample size and monitoring technique that alarms when the fraction defective increases to an unacceptably high level, but does not alarm when the process is at nominal. The nominal (acceptable) fraction defective is used, and an increased fraction defective (unacceptable) is assumed as part of the control chart design. The control chart recommended for this problem is the Bernoulli Cumulative Sum (CUSUM) control chart.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

This page left blank

This page left blank

ACRONYMS AND TERMS

Acronym/Term	Definition
ARL	Average Run Length
BC	Bernoulli CUSUM
HVB	High Voltage Breakdown
MRL	Median Run Length
NSE	Nuclear Security Enterprise

1. INTRODUCTION

The primary goal of a shelf life plan is to quickly detect a change that would threaten the component's ability to deliver the desired output within specifications over its required stockpile life. In this report a statistical framework is outlined for establishing a shelf-life program for components whose performance is measured by a binary response variable, usually 'pass' or 'fail.' The approach applies to both single measurement devices and repeated measurement devices, although additional process control charts may be useful in the case of repeated measurements. Shelf life plans for monitoring variables data were previously developed in Crowder (2017).

A type of deviation from nominal that will be modeled in this report is a **step-change or gradual increase** in fraction defective. This is a fraction defective that starts at nominal or better but instantaneously or gradually increases to an unacceptably high level as the component ages. Possible causes include wear out, corrosion, or other aging phenomena that may induce a catastrophic failure. A control chart of shelf life data should produce an alarm so that this condition can be mitigated before the component threatens a significant portion of the stockpile. An example of a gradual change model is in Figure 1 below.

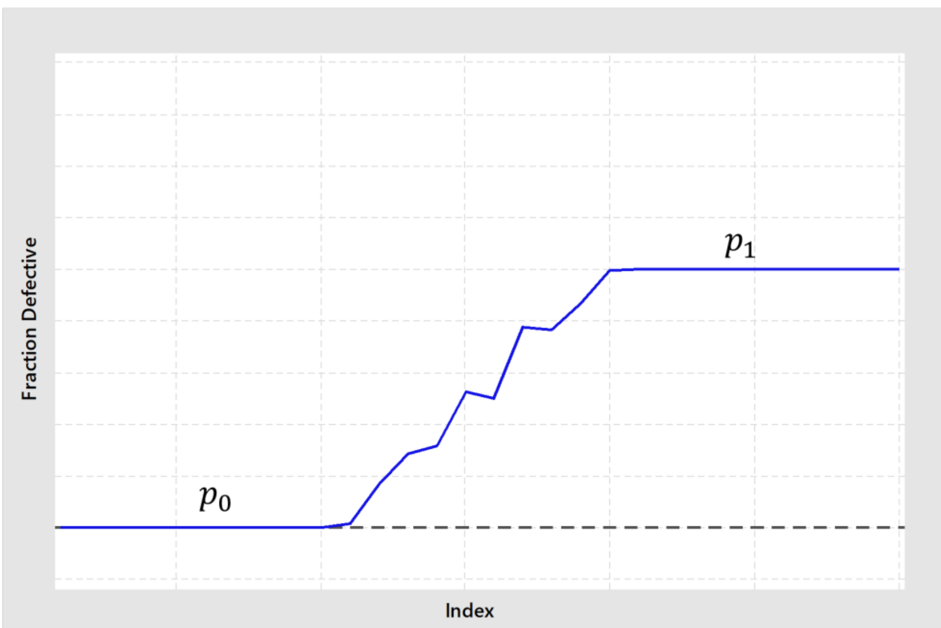

Figure 1. Gradual Increase in Fraction Defective

In this figure p_0 is the nominal fraction defective and p_1 is the minimum unacceptable fraction defective.

The report is divided into three additional sections. An overview of the Bernoulli CUSUM (BC) is presented first, along with a simulated example of its use. In the following section, general recommendations are made for designing a BC control chart to monitor fraction defective, assuming a step-change or gradual increase in fraction defective. Properties of the BC chart in terms of its run length distribution are presented as part of the design. Finally, recommendations are made (Appendix B) for designing a BC chart to monitor data from a shelf-life program at Sandia Labs.

2. AN OVERVIEW OF THE BERNOULLI CUSUM

In probability and statistics, a *Bernoulli Process* is a sequence of independent binary random variables $X_1, X_2, X_3, \ldots$, that each take on the value 0 or 1. The random variable X_t takes on the value 1 with probability p and takes on the value 0 with probability $(1 - p)$. In manufacturing, we can think of a sequence of manufactured parts as being assigned the value 1 if the part is defective, and the value 0 if the part functions properly. Then p represents the manufactured fraction defective. It is of interest to monitor the fraction defective and provide timely feedback to process engineers if the fraction defective is believed to have increased.

The BC control chart is a statistical process monitoring technique used to detect changes in the fraction defective p, from a nominal value p_0 to an unacceptable level p_1. It "cumulates" the number of defects that occur in a manufacturing window and provides an ongoing test of whether the fraction defective has increased. The Bernoulli CUSUM chart has appeared in the recent statistical process control literature (see Reynolds and Stoumbos (1999), Szarka (2011), Szarka and Woodall(2011), and Crowder (2017)), used primarily for high quality, high volume processes. A challenge has been to use the chart for high quality processes with somewhat lower volume. Several control charts, including the traditional p-chart, can be used to monitor processes with binary data. The p-chart monitors the fraction defective in successive samples, with a minimum recommendation of 25 to 50 parts per sample. Other control charts suggested for this problem include the Binomial CUSUM, applied to the number of failures per sample, and the Geometric CUSUM, applied to the number of good parts between failures.

A primary advantage of the BC is that the plotted statistic is calculated after the inspection of *each* part. Because of this property, it has the best statistical properties in terms of detecting increases in fraction defective for high quality processes (Szarka, 2011). By "best statistical properties" it is meant that this type of control chart will detect increases in fraction defective more quickly than competing control charts.

The upper one-sided Bernoulli CUSUM statistics, $B_t, t = 1, 2, 3, \ldots,$ are

$$B_t = max\,(0, B_{t-1} + \ (X_t - r)), \tag{1}$$

where $B_0 = 0$ and r is a small constant greater than zero but less than one. The X_t's represent the random Bernoulli sequence of 0's and 1's. The BC produces an alarm if $B_t \geq H$, a threshold value that is chosen, along with r, as part of the CUSUM design. The Likelihood Ratio Test for testing a simple hypothesis of p_0 vs. p_1 leads to the Bernoulli CUSUM statistic as optimal. Reynolds and Stoumbos (1999) discuss this relationship and recommend choice of r.

The measure of performance often used to evaluate the Bernoulli CUSUM is the Average Run Length (ARL), defined as the expected number of parts tested until the threshold H is exceeded. A signal on the BC chart provides a warning that the process fraction defective may have increased. An investigation of the process would follow any such alarm. Because the Run Length distribution is highly skewed, we will instead use the Median Run Length (MRL) as the primary measure of performance. This same measure helps determine the best possible design for the CUSUM chart regarding choice of H and r.

An example of how the Bernoulli CUSUM works is given in the Figure 2 below.

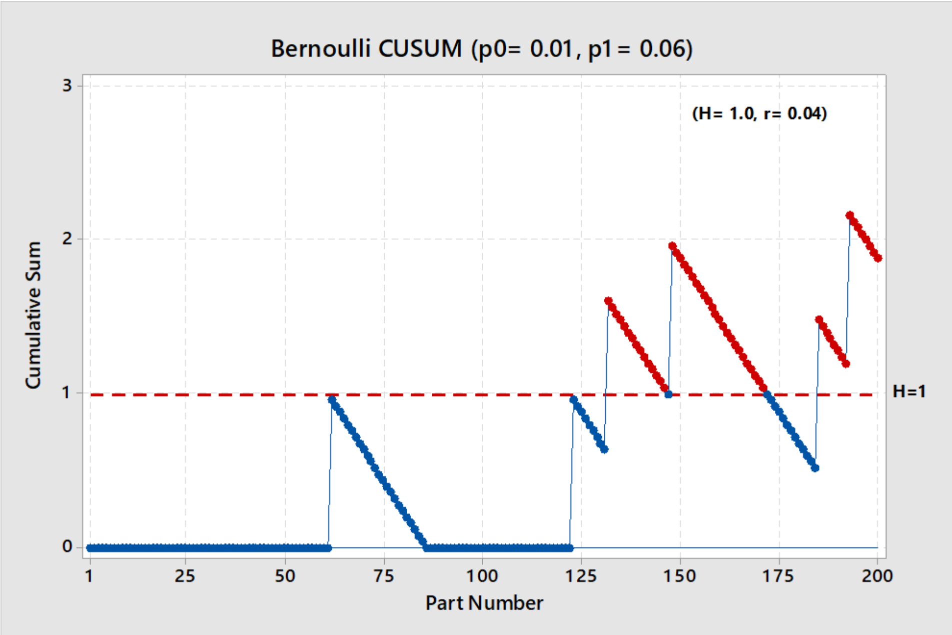

Figure 2 – Example Bernoulli CUSUM Control Chart with (H, r)=$(1.0, 0.04)$.

This example uses simulated data with an initial defect rate of $p_0 = 0.01$ for the first 100 parts followed by a defect rate of $p_1 = 0.06$ for the second 100 parts. The control limit is $H = 1.0$ and the reference value is $r = 0.04$. The plotted Bernoulli CUSUM value is zero until the first defect occurs at part number 62, where the CUSUM increases to the value $(1 - 0.04) = 0.96$ (applying Equation 1). From that point forward, the CUSUM decreases by 0.04 for each part that passes until it returns to zero. With two failures at part numbers 123 and 132, the CUSUM signals because two failures have occurred in a relatively small window of parts. With H =1.0 and r =0.04, the CUSUM will signal whenever 2 failures occur within a window of $1/r$ = 25 parts. In Figure 2, the BC chart has been plotted beyond the signal at observation 132. In practice, this alarm would be reported to the surveillance team for further investigation.

Advantages of the Bernoulli CUSUM compared to other control charts used for binary data:

1. The method has been shown to detect increases in the process fraction defective faster than competing methods, measured by Median Run Length. It is used to answer the question: Has the fraction defective increased?
2. The method has the advantage of testing for an increase in fraction defective after *each* part. There is no need to accumulate parts before testing for an increase.
3. The method provides a “moving window” of current process performance.
4. The method applies to process manufacturing data, product acceptance data, and shelf-life data. The ordering of the individual data values must, of course, be meaningful.
5. The method is relatively easy to explain and implement, and standard statistical packages can be used to plot the Bernoulli CUSUM.

The advantages listed above have led to the development of the Bernoulli CUSUM for monitoring the production of high reliability, high consequence electrical components. We discuss this application below. The goal is that the BC will provide an early indication of an increase in fraction defective during product acceptance testing or shelf-life testing. In the next section a strategy is presented for design of a Bernoulli CUSUM control chart.

3. DESIGN OF THE BERNOULLI CUSUM

The measure of performance often used to evaluate the Bernoulli CUSUM is the Average Run Length (ARL), defined as the expected number of parts tested until the threshold H is exceeded. A signal on the BC chart provides an alarm that the process fraction defective may have increased. An investigation of the process would follow any such alarm.

Because the Run Length distribution is highly skewed, however, we will instead use the Median Run Length (MRL) as the primary measure of performance. This measure is used to determine the "best" design for the CUSUM chart in terms of choice of H and r. Given a nominal value p_0 and an unacceptable value p_1, the recommended design of the Bernoulli CUSUM is given by the following steps:

1. As a preliminary choice for r, use the expression

$$r = \frac{-log(\frac{1-p_1}{1-p_0})}{log(\frac{p_1(1-p_0)}{p_0(1-p_1)})}. \tag{2}$$

This value is recommended by Reynolds and Stoumbos (1999). This choice is based on the representation of the CUSUM chart as the optimal outcome of a sequential probability ratio test (SPRT), testing p_0 vs. p_1. The reference value must satisfy $0.0 < r < 1.0$, but it will typically be close to zero.

2. Choose the control limit H. This value affects both the false alarm rate and the number of parts tested until an increase in fraction defective is detected. A reasonable approach is to choose H so that the Median Run Length (MRL) is large enough to avoid any false alarms over the life of use of the BC when the fraction defective remains at nominal $(p = p_0)$.

3. For the choice of (H, r) determined from Steps 1 and 2, evaluate the MRL for values of the fraction defective that are greater than nominal $(p > p_0)$. The MRL when the fraction defective exceeds nominal is the "time to detection" of an unacceptable fraction defective. Additional choices of (H, r) can easily be explored via simulation or by using tabled values of MRLs (see Appendix A).These tables show sensitivity to small changes in both H and r.

Tables of MRLs in Appendix A were generated via simulation for values of H and r with ranges $1.0 \le H \le 3.0$ and $0.01 \le r \le 0.04$. These tables can be used to identify a starting point for (H, r) using the steps above. Percentiles of the Run Length Distribution can used for a more detailed analysis of the BC performance, and to make probability statements regarding possible outcomes.

An example of how to use these tables in a sensitivity analysis is given next. In this example the values of H being compared are $H = 1.0$ and $H = 1.4$. The vales of the reference r being compared are $r = 0.01$ and $r = 0.04$.

Table 1. Median Run Lengths for various (H, r) combinations

p	H=1.0 r=0.01	H=1.4 r=0.01	H=1.0 r=0.04	H=1.4 r=0.04
0.01	187	227	413	605
0.02	84	89	132	177
0.03	56	56	72	90
0.04	43	42	49	58
0.05	34	33	35	42
0.06	28	28	28	34
0.07	24	24	24	27
0.08	21	21	21	22
0.09	19	18	19	20
0.10	17	17	17	18

This table shows that the MRLs increase as H increases with r fixed, for the smallest values of p, and decrease as r decreases with H fixed, for $p \leq 0.05$. The control limit H can be adjusted to produce a desired MRL for the nominal case, and the value r can be used to find the minimum MRL for a given value of p greater than nominal. These tables can be supplemented with additional percentiles of the run length distribution.

For the special case (H, r) = (1.0, 0.04), the full run length distributions (with 5th, 50th, and 95th percentiles identified) appear below for p = 0.01 (Figure 3) and p = 0.06 (Figure 4).

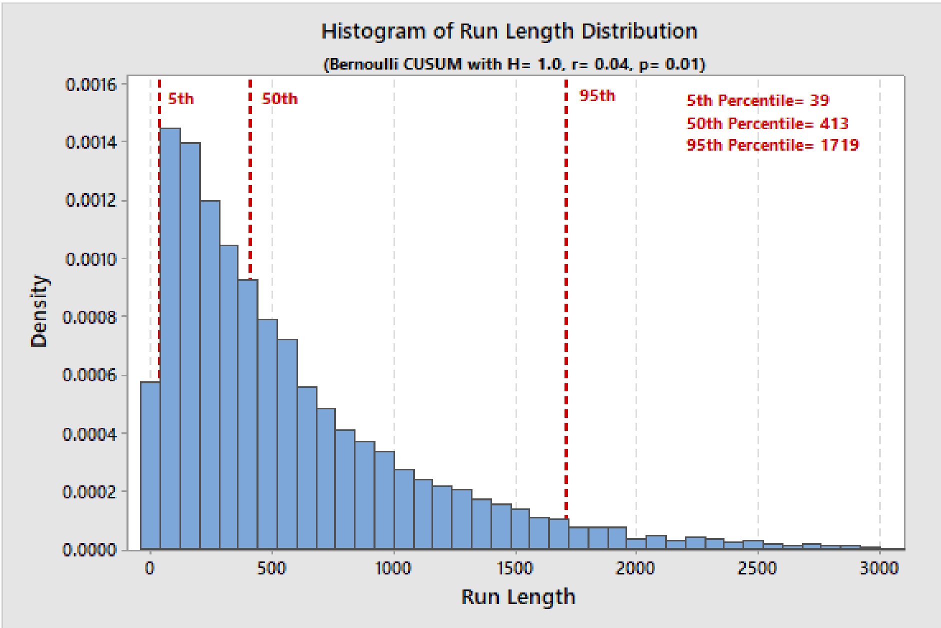

Figure 3 - Run Length Distribution with p = 0.01.

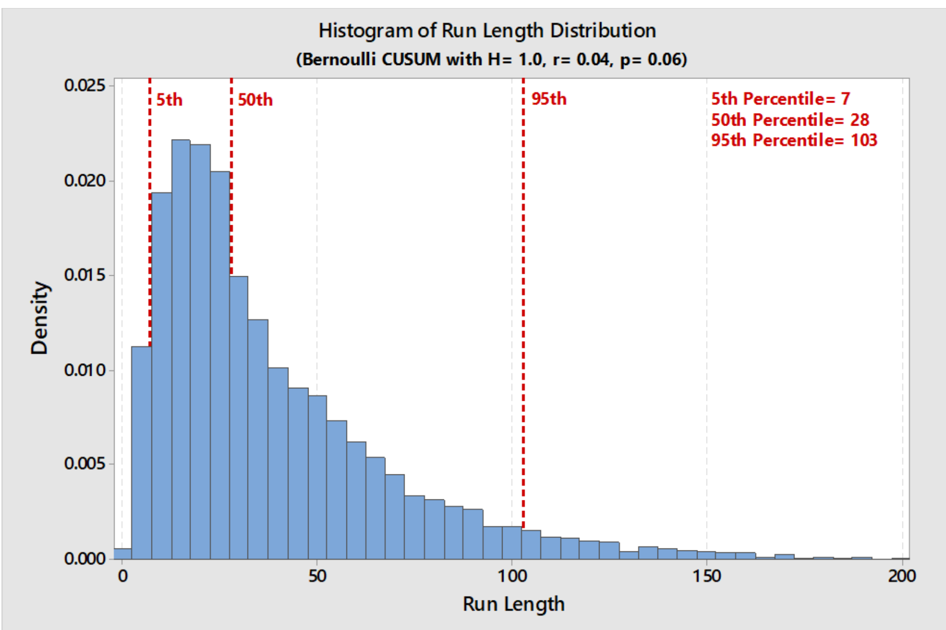

Figure 4 - Run Length Distribution with p = 0.06.

These histograms, each based on 10,000 simulations, show skewness in the run length distributions for p = 0.01 and p = 0.06. In these cases, the 50th percentile provides the best estimate of central tendency, and the 5th and 95th percentiles provide a 90% probability interval for Run Length outcomes. For the Run Length distribution with p = 0.01, these values (Figure 3) are interpreted to mean that 50% of the time the chart will signal within approximately 410 observations, and 95% of the time the chart will signal within approximately 1700 observations. This can be interpreted as the false alarm rate. For the Run Length distribution with p = 0.06, these values (Figure 4) are interpreted to mean that 50% of the time the chart will signal within approximately 30 observations, and 95% of the time the chart will signal within approximately 100 observations. This can be interpreted as the time (number of parts) to detection when the fraction defective is unacceptably high.

4. BERNOULLI CUSUM TO MONITOR PRODUCTION PERFORMANCE

The Bernoulli CUSUM was used to monitor the production of a high-reliability, high-consequence electronic part manufactured within the Nuclear Security Enterprise (NSE). The design of an appropriate BC chart for this problem, following the steps above, is described in this section.

The part requires some manual operator assembly, so a nominal value of p_0= 0.005 is considered the lowest reasonably attainable fraction defective. Production personnel perform 100% inspection and the most common failure mode is high voltage breakdown (HVB). A single part is very expensive, so timely feedback regarding any increase in fraction defective is critical. False alarms are also costly, so a median run length when p= 0.005 is desired to be at least 8000 parts. It is desired that the BC control signals quickly whenever the fraction defective is at or exceeds p_1= 0.05.

Step 1. The preliminary choice for r is (using expression (2))

$$r = \frac{-log(\frac{1-0.05}{1-0.005})}{log(\frac{0.05(1-0.005)}{0.005(1-0.05)})}.$$

$$= \frac{-log(\frac{0.95}{0.995})}{log(\frac{0.05(0.995)}{0.005(0.95)})}$$

$$\cong 0.02.$$

Step 2. The proposed Bernoulli CUSUM design strategy results in the following condition:

Subject to MRL ≥ 8000 when p_0= 0.005,

choose the best overall combination of (H, r) when p_1= 0.05.

When the process is operating at fraction defective p_0 or lower, it is desirable to have a large MRL, to minimize false alarms. The median run length when p_0= 0.005 is therefore required to be 8000 or greater. The value p_1 is the minimum unacceptable fraction defective and must be detected quickly. When the process is operating at this level or greater, it is desirable to have a small MRL.

Expression (2) was used to provide the starting reference value $r = 0.02$. Monte Carlo simulation with searches were then used to identify various combinations of (H, r) for $0.005 \leq p \leq 0.05$ that produce an MRL of approximately 8000 when p_0= 0.005 (Table 2). It should be noted that the MRL values that appear in Table 2 are approximate, each based on 10,000 simulations.

Table 2. Combinations of (H, r) that Produce a Nominal MRL of Approximately 8000.

p	H =2.2 r = 0.020	H =2.4 r = 0.017	H =2.6 r = 0.014	H =2.8 r = 0.012	H =3.0 r = 0.0105
0.005	8000	8000	8000	8000	8000
0.01	1274	1153	1013	971	881
0.02	255	244	232	233	231
0.03	123	123	123	130	134
0.04	80	83	85	89	93
0.05	58	62	66	71	74

From this table we can see that MRLs greatly vary for each combination of (H, r) when the fraction defective is greater than $p_0 = 0.005$. Larger H values result in quicker detection when p = 0.01 or 0.02, but slower detection when p = 0.03, 0.04 or 0.05. Since we are more interested in quicker detection when p = 0.01 or 0.02, the recommended choice is to use (H, r) = (3.0, 0.0105). This combination of (H, r) will produce a signal on the Bernoulli CUSUM chart if there are 4 or more failures within $1/0.0105 \cong$ 95 or fewer parts. Table 3 gives the 5th, 25th, 50th, 75th and 95th percentiles of the Run Length distribution for this BC chart.

Table 3. Percentiles of Run Length Distribution of Bernoulli CUSUM with (H, r) = (3.0, 0.0105)

p	5th	25th	50th	75th	95th
0.005	692	3259	8000	15585	30455
0.01	172	463	881	1540	3500
0.02	68	139	231	375	645
0.03	45	83	134	203	353
0.04	35	63	93	138	236
0.05	27	50	74	100	178

In the row of Table 3 with fraction defective p= 0.01, the MRL (50th percentile) is 881, the 5th percentile is 172 and the 95th percentile is 3500. These extreme percentiles provide a "best case" and "worst case" number of parts that will be needed to detect an increase in fraction defective to p= 0.01. To lower these numbers, the MRL when p= 0.005 would also have to be lowered, resulting in an increased false alarm rate. The choice of (H, r) = (3.0, 0.0105) is an attempt to balance the desire to quickly detect an increase in fraction defective with the desire to keep the false alarm rate low.

A retrospective analysis of electronic part pass/fail data was performed using this Bernoulli CUSUM design. The plot of the resulting BC chart appears below.

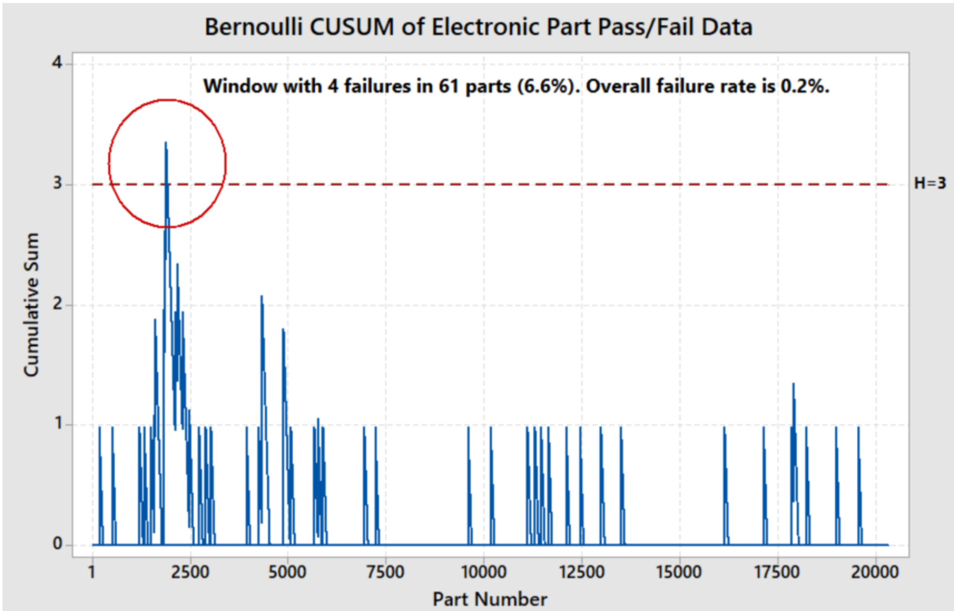

Figure 5 - Bernoulli CUSUM of Electronic Part Pass/Fail Data

This analysis suggests a process problem occurred around part number 1800 when 4 failures occurred within a window of 61 parts (fraction defective is $p = 0.066$) The analysis also indicated, however, that the problem did not persist, with an overall fraction defective of $p = 0.002$, well below the nominal target. The Bernoulli CUSUM chart with these parameters was implemented to monitor ongoing production and is presently in use.

Appendix A below presents tables of Median Run Lengths for various choices of Bernoulli CUSUM parameters (H, r). Appendix B outlines the design of a Bernoulli CUSUM chart to use in a shelf-life program when only binary data are available.

REFERENCES

Bourke, Patrick (1991). Detecting a Shift in Fraction Nonconforming Using Run-Length Control Charts With 100% Inspection. *Journal of Quality Technology*, Vol. 23, No. 3, pp 225-238.

Bourke, Patrick (2001). Sample Size and the Binomial CUSUM Control Chart: The Case of 100% Inspection. *Metrika*, 53: 51-70.

Crowder, Stephen (2017). An Introduction to the Bernoulli CUSUM. RAMS Conference 2017.

Lucas, James (1985). Counted Data CUSUM's. *Technometrics*, Vol. 27, No. 2, pp. 129-144.

Lucas, James (1989). Control Schemes for Low Count Levels. *Journal of Quality Technology*, Vol. 21, No. 2, pp 199-201.

Reynolds, Marion and Zachary Stoumbos (1999). A CUSUM Chart for Monitoring a Proportion When Inspecting Continuously. *Journal of Quality Technology*, Vol. 31, No. 1, pp 87-108.

Sego, Landon (2006). Applications of Control Charts in Medicine and Epidemiology. Unpublished Dissertation, Virginia Polytechnic Institute and State University. (Available Online)

Szarka, John (2011). Surveillance of Negative Binomial and Bernoulli Processes. Unpublished Dissertation, Virginia Polytechnic Institute and State University. (Available Online)

Szarka, John and William Woodall (2011). A Review and Perspective on Surveillance of Bernoulli Processes. *Quality and Reliability Engineering International*, Vol. 27, No. 6, pp 735-752.

APPENDIX A. TABLES OF MRLS FOR THE BERNOULLI CUSUM CHART

The range of values in these tables should be sufficient to design a Bernoulli CUSUM chart for shelf-life programs typically used within the NSE. If additional combinations of (H, r) need to be investigated, the user can do so via simulation using the author's Matlab script (Appendix C) that is available.

Table A1. Median Run Lengths for $H = 1.0$, 1.2, 1.4, 1.6, 1.8, and 2.0 with $r = 0.01$.

p	$H = 1.0$ $r = 0.01$	$H = 1.2$ $r = 0.01$	$H = 1.4$ $r = 0.01$	$H = 1.6$ $r = 0.01$	$H = 1.8$ $r = 0.01$	$H = 2.0$ $r = 0.01$
0.01	188	203	225	269	328	428
0.02	83	84	89	102	119	145
0.03	56	56	56	60	74	89
0.04	42	42	42	42	52	66
0.05	34	34	34	34	39	54
0.06	28	28	28	28	30	45
0.07	24	24	24	24	25	38
0.08	21	21	21	21	21	33
0.09	19	19	19	19	19	30
0.10	17	17	17	17	17	27

Table A2. Median Run Lengths for $H = 1.0$, 1.2, 1.4, 1.6, 1.8, and 2.0 with $r = 0.02$.

p	$H = 1.0$ $r = 0.02$	$H = 1.2$ $r = 0.02$	$H = 1.4$ $r = 0.02$	$H = 1.6$ $r = 0.02$	$H = 1.8$ $r = 0.02$	$H = 2.0$ $r = 0.02$
0.01	257	294	334	429	614	966
0.02	98	104	115	137	170	218
0.03	56	59	66	75	92	107
0.04	42	43	45	52	60	74
0.05	34	34	34	39	46	54
0.06	28	29	28	31	39	45
0.07	24	25	24	25	31	39
0.08	21	21	21	21	26	33
0.09	19	19	19	19	23	30
0.10	17	17	17	17	20	27

Table A3. Median Run Lengths for $H = 1.0$, 1.2, 1.4, 1.6, 1.8, and 2.0 with $r = 0.03$.

p	$H = 1.0$ $r = 0.03$	$H = 1.2$ $r = 0.03$	$H = 1.4$ $r = 0.03$	$H = 1.6$ $r = 0.03$	$H = 1.8$ $r = 0.03$	$H = 2.0$ $r = 0.03$
0.01	327	391	474	618	979	1870
0.02	114	121	142	175	238	342
0.03	62	68	79	91	113	145
0.04	43	45	51	59	72	87
0.05	33	35	37	43	52	62
0.06	28	27	31	35	42	49
0.07	24	24	25	28	35	39
0.08	21	21	21	24	29	34
0.09	19	19	19	21	26	31
0.10	17	17	17	17	23	27

Table A4. Median Run Lengths for $H = 1.0$, 1.2, 1.4, 1.6, 1.8, and 2.0 with $r = 0.04$.

p	$H = 1.0$ $r = 0.04$	$H = 1.2$ $r = 0.04$	$H = 1.4$ $r = 0.04$	$H = 1.6$ $r = 0.04$	$H = 1.8$ $r = 0.04$	$H = 2.0$ $r = 0.04$
0.01	395	490	625	850	1500	3300
0.02	133	150	178	229	334	515
0.03	70	79	92	114	149	203
0.04	47	51	59	70	88	111
0.05	37	40	42	49	61	74
0.06	29	30	33	37	46	54
0.07	24	25	27	32	38	44
0.08	21	21	23	26	32	37
0.09	18	19	20	23	27	32
0.10	17	17	18	20	24	28

Table A5. Median Run Lengths for $H = 1.0$, 1.2, 1.4, 1.6, 1.8, and 2.0 with $r = 0.05$.

p	$H = 1.0$ $r = 0.05$	$H = 1.2$ $r = 0.05$	$H = 1.4$ $r = 0.05$	$H = 1.6$ $r = 0.05$	$H = 1.8$ $r = 0.05$	$H = 2.0$ $r = 0.05$
0.01	475	610	755	1125	2200	5000
0.02	147	172	211	283	465	745
0.03	81	95	107	136	196	263
0.04	55	57	67	82	104	138
0.05	38	43	48	57	75	89
0.06	29	33	36	44	54	64
0.07	25	27	31	35	42	51
0.08	21	22	25	29	35	41
0.09	19	19	22	25	30	35
0.10	17	17	18	22	25	31

In Appendix B, an example is presented to illustrate the design of a shelf-life program using the values tabled above.

APPENDIX B. DESIGN OF A BERNOULLI CUSUM FOR A SHELF-LIFE PROGRAM

Suppose that a shelf-life plan is to be developed for a component that is required to be in the stockpile for twenty years. Suppose further that the component's reliability requirement is $R = 0.99$ ($p_0 = 0.01$), and the minimum acceptable reliability is $R = 0.95$ ($p_1 = 0.05$). The Bernoulli CUSUM chart should thus be designed to rarely signal when the fraction defective is $p = 0.01$, but quickly signal when the fraction defective increases to $p \geq 0.05$.

The Median Run Length (MRL) will again be used as the primary measure of performance, and it will be used to determine the best BC chart design in terms of (H, r). The design of the Bernoulli CUSUM follows the steps outlined above.

1. As a preliminary choice, set the control limit reference value equal to

$$r = \frac{-log(\frac{1-p_1}{1-p_0})}{log(\frac{p_1(1-p_0)}{p_0(1-p_1)})}. \qquad \text{(B1)}$$

Substituting $p_0 = 0.01$ and $p_1 = 0.05$ into (B1) results in $r \cong 0.025$. The tables in Appendix A do not include $r = 0.025$ as an exact choice, so the focus is instead on $r = 0.02$ and $r = 0.03$.

2. Choose the control limit H. This value helps determine the false alarm rate, the number of parts that are tested until the BC signals under nominal conditions. A reasonable approach is to choose the nominal MRL to be larger than the total number of parts that could be tested during the entire shelf life program. The following table was constructed from subsets of tables in Appendix A, using $r = 0.02$ and 0.03 with $H = 1.0, 1.2,$ and 1.4.

Table B1. Possible Choices of (H, r) for Bernoulli CUSUM chart.

p	$H = 1.0$ $r = 0.02$	$H = 1.2$ $r = 0.02$	$H = 1.4$ $r = 0.02$	$H = 1.0$ $r = 0.03$	$H = 1.2$ $r = 0.03$	$H = 1.4$ $r = 0.03$
0.01	257	294	334	327	391	474
0.02	98	104	115	114	121	142
0.03	56	59	66	62	68	79
0.04	42	43	45	43	45	51
0.05	34	34	34	33	35	37
0.06	28	29	28	28	27	31
0.07	24	25	24	24	24	25
0.08	21	21	21	21	21	21
0.09	19	19	19	19	19	19
0.10	17	17	17	17	17	17

3. For the choices of (H, r) determined from Steps 1 and 2, compare the MRLs for values of p that are greater than nominal. These are the median number of "parts tested until detection" of an

increase in fraction defective. Table B1 shows that the number of parts to detection is approximately the same for all choices in the table whenever $p \geq 0.05$. The choice $(H, r) = (1.0, 0.02)$ has the quickest detection of any choice, but it also has the smallest MRL when $p = 0.01$. This means it has the highest likelihood of a false alarm over the life of the program. A reasonable compromise between detection and false alarm rates is $(H, r) = (1.2, 0.03)$. Percentiles of the Run Length distribution could also be used for a more detailed comparison of the options in the table. If it is decided that $(H, r) = (1.2, 0.03)$ gives acceptable performance for all values of $0.01 \leq p \leq 0.10$, it would be taken as the final choice.

The number of parts to test should then be based on how quickly the increase in fraction defective must be detected. Table B1 shows that the median number of measurements to detect an increase in fraction defective from $p = 0.01$ to $p = 0.05$ is approximately 35. If detection within a few years is necessary, then the same 35 parts would be tested annually. The median number of measurements to detect an increase in fraction defective from $p = 0.01$ to $p = 0.10$ is 17. If detection within a few years is necessary, then the same 17 parts would be tested annually. Choice of sample size thus becomes a compromise between cost of parts plus testing and cost of delaying detection. This number could be rounded to 20 to be close to the so-called "90/10" stockpile surveillance sampling number of 22 parts. **A rule of thumb recommendation is thus to test 20 parts each year to quickly detect a 10% or worse problem**.

An implicit assumption is that the probability of failure does not increase simply as the result of testing. The type of failure mechanisms that would be detected with this approach are those associated with manufacturing defects that are precipitated by aging, but not by repeated testing.

Several cautions must be made, however, with respect to control charting with repeated measurements from a limited number of samples.

1. Process changes made during production can result in changes in performance and the appearance of sub-populations. A small sample of components measured repeatedly over time is less likely to be representative of the entire population and such changes may not be reflected in the shelf life parts.
2. Repeated measurements may degrade the performance of the component over time.
3. The small sample size cannot accommodate the potential need for experimental parts outside the shelf life study.
4. Fewer parts will be available for reliability estimation. This problem may be compounded by the desire to reduce testing quantities in the future.

The parts taken for the shelf life program should be collected during each production year to make them as representative as possible.

For single measurement devices, the total sample size would be the number chosen using Steps 1-3 above, times the number of years allotted to testing. For example, if testing in a 30-year program is to be done on years 5, 10, 15, 20, and 25, and 30, the rule of thumb requirement would be (6 x 20)= 120 parts to quickly detect $p_1 \geq 0.10$ and (6 x 35)= 210 parts to quickly detect $p_1 \geq 0.05$.

Table B2 below is a simulated outcome of testing the same 20 parts each year, when the underlying fraction defective remains constant at $p = 0.01$. It is assumed that a failed unit is not re-tested. Failures are denoted by the red number '1,' and successes are denoted by '0.'

Table B2. Simulated Outcome of Twenty Parts Each Tested Annually for Twenty Years

Yr\Part	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	**1**	0	0	0	0		0	0	0	0	0	0	0	0	0
9	0	0	0	0	0		0	0	0	0		0	0	0	0	0	0	0	0	0
10	0	0	0	0	0		0	0	0	0		0	0	0	0	0	0	0	0	0
11	0	0	0	0	0		0	0	0	0		0	0	0	0	0	0	0	0	0
12	0	0	0	0	0		0	0	0	0		0	0	0	0	0	0	0	0	0
13	0	0	0	0	0		0	0	0	0		0	0	0	0	0	0	0	0	0
14	0	**1**	0	0	0		0	0	0	0		0	0	0	0	0	0	0	0	0
15	0		0	0	0		0	0	0	0		0	0	0	0	0	0	0	0	0
16	0		0	0	0		0	0	0	0		0	0	0	0	0	0	0	0	0
17	0		0	0	0		0	0	0	0		0	0	0	0	0	0	0	0	0
18	0		0	0	0		0	0	0	0		0	0	0	0	0	**1**	0	0	0
19	0		0	0	0		0	0	0	0		0	0	0	0	0		**1**	0	0
20	0		0	0	0		0	0	0	0		0	0	0	0	0			0	0

The resulting BC Chart, with $(H, r) = (1.2, 0.03)$ appears below. Note from Table B1 that the "false alarm" rate is 1 in 390. The false alarm in this chart appears at part number 334, consistent with the estimated median.

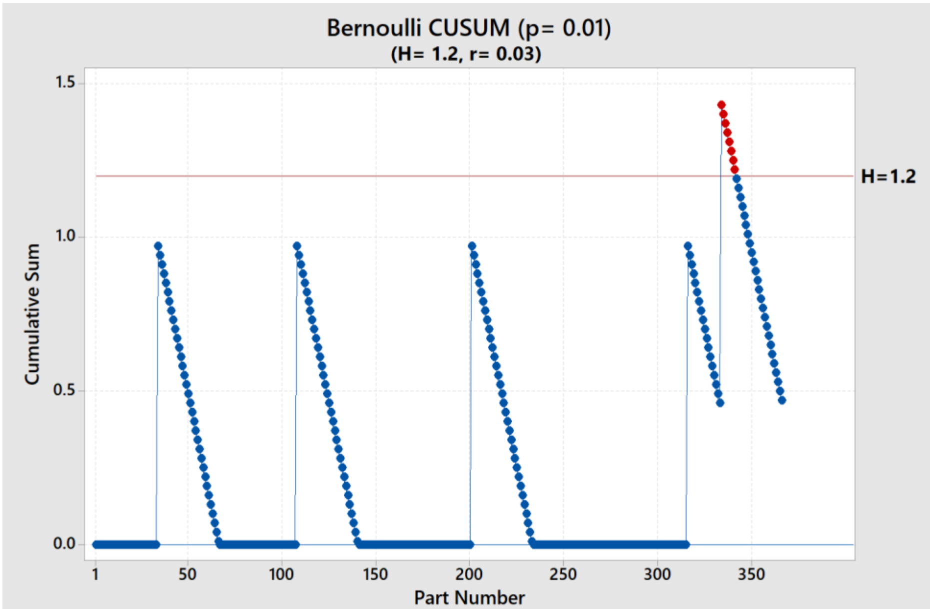

Figure B1. Bernoulli CUSUM Chart of Simulated Part Data with $(H, r) = (1.2, 0.03)$.

APPENDIX C. MATLAB SCRIPT FOR EVALUATING RUN LENGTH PERCENTILES

```
%
% This program evaluates the run length distribution for the Bernoulli
% CUSUM Control Chart used to detect changes in fraction defective.
%
clear all;
% Reference value (set to ref=0.02)
ref=0.02;
% Upper Control Limit (H) (set to UCL= 2.0)
UCL=2.0;
fprintf('%4.2f %5.4f \n',UCL,ref);
% Set the number of Simulated Runs (set to b=10,000)
b=10000;
% Enter the fraction defective (fd) values (0.01 to 0.10 by 0.01)
fd(1)=0.01;fd(2)=0.02;fd(3)=0.03;fd(4)=0.04;fd(5)=0.05;
fd(6)=0.06;fd(7)=0.07;fd(8)=0.08;fd(9)=0.09;fd(10)=0.10;
%
for i1=1:10
 p=fd(i1);
for i=1:b
 % Initialize B_CUSUM (set to 0.0); Value chosen is a "head start"
 B_CUSUM=0.0;
 K=1;
 while B_CUSUM<UCL
 % Compute next value of the CUSUM
    X=binornd(1,p);
    B_CUSUM=max(0.0,B_CUSUM +(X-ref));
    K=K+1;
 end
 Run(i)=K-1;
end
% Compute the Average, 5th, 25th, 50th, 75th, and 95th percentiles
% of the Run Length Distribution
RL_Ave=mean(Run);
RL_5th=prctile(Run,5);
RL_25th=prctile(Run,25);
RL_50th=prctile(Run,50);
RL_75th=prctile(Run,75);
RL_95th=prctile(Run,95);
run_length(i1,1)=p;run_length(i1,2)=RL_Ave;run_length(i1,3)=RL_5th;
run_length(i1,4)=RL_25th;run_length(i1,5)=RL_50th;
run_length(i1,6)=RL_75th;run_length(i1,7)=RL_95th;
% Print the Average, 5th, 25th, 50th, 75th, and 95th percentiles
% of the Run Length Distribution
fprintf('%4.3f %8.2f %8.2f %8.2f %8.2f %8.2f %8.2f \n',p,RL_Ave,RL_5th,...
RL_25th,RL_50th,RL_75th,RL_95th);
end
```

Output from above Matlab script:

```
2.00 0.0200
0.010  1379.23   114.00   428.00   960.50  1912.00  4048.50
0.020   292.56    42.00   107.00   214.50   394.00   808.00
0.030   143.92    28.00    60.00   108.00   191.00   383.50
0.040    92.89    21.00    43.00    73.00   121.00   233.50
0.050    68.44    18.00    35.00    55.00    89.00   165.00
0.060    54.33    15.00    29.00    44.00    70.00   127.00
0.070    44.85    13.00    25.00    38.00    56.00   101.00
0.080    38.66    11.00    22.00    33.00    48.00    86.00
0.090    34.23    10.00    20.00    30.00    43.00    74.00
0.100    30.31     9.00    18.00    27.00    39.00    63.00
```

The first line of output gives the values of H and r. Successive lines of the output give the values of p, ARL, 5th, 25th, 50th, 75th, and 95th percentiles of the run length distribution.

DISTRIBUTION

Email—Internal

Name	Org.	Sandia Email Address
Matthew D. Smith	05574	mdsmith@sandia.gov
Staff of Department 05574	05574	ORG-05574@sandia.gov
Technical Library	01911	sanddocs@sandia.gov

This page left blank

This page left blank

Sandia National Laboratories

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.