

# Reliability Growth Testing

## Introduction

The objective of this lab is to gain familiarity with Software Quality and Reliability Assessment Tools. This lab is composed of two parts; the first's focus being on Reliability Growth Testing and the second's focus being on RDC Demo Charts.

## Result of Model Comparison

*(Top Two Models)*

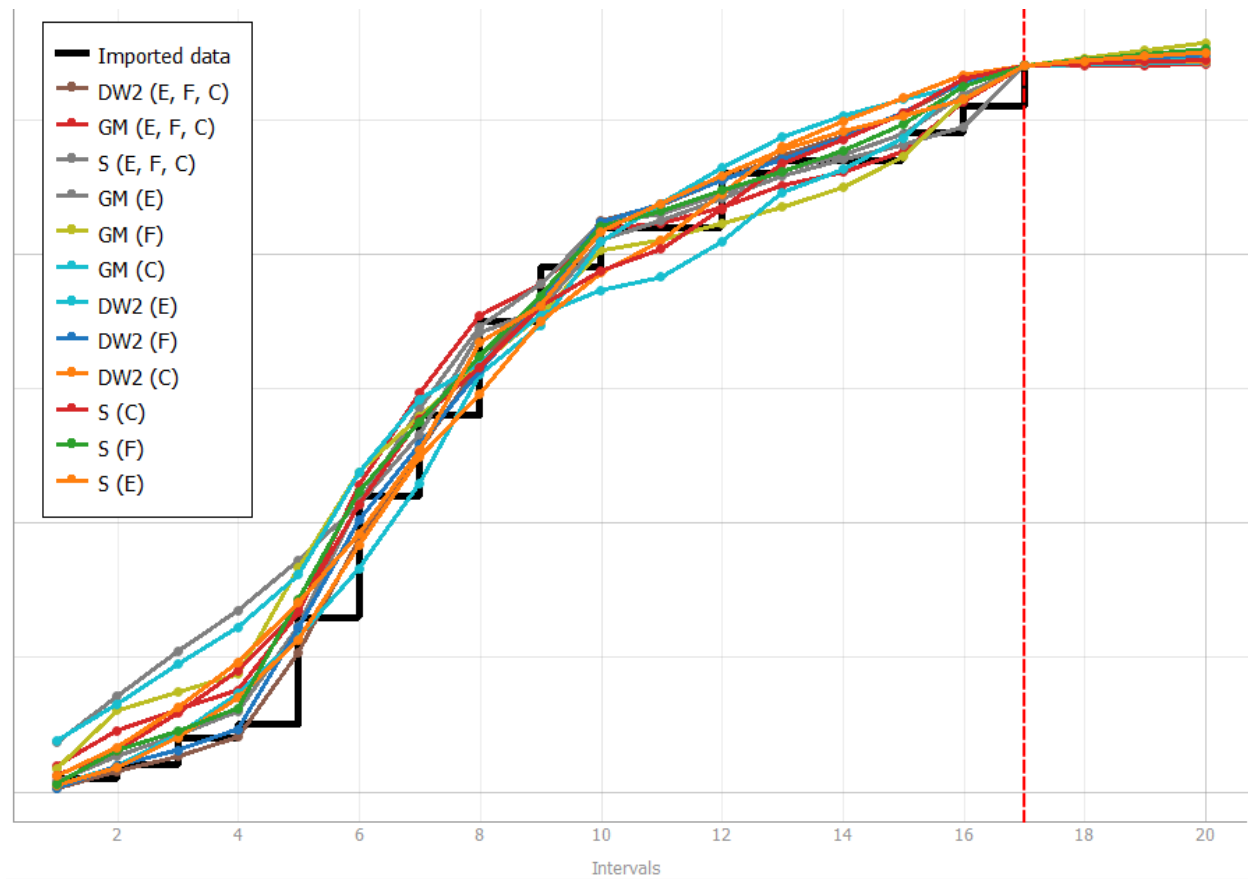
	Model Name	Covariates	Log-Likelihood	AIC	BIC	SSE	PSSE	Critic (Mean)	Critic (Median)
2	S	E	-32.003	72.005	75.338	65.239	13.492	0.571	0.510
3	S	F	-28.718	65.436	68.769	23.004	28.929	1.000	1.000
4	S	C	-31.678	71.357	74.690	70.462	18.678	0.601	0.550
8	S	E, F, C	-27.137	66.275	71.274	9.178	2.875	0.997	1.000
10	DW2	E	-33.200	72.400	74.900	122.824	16.403	0.459	0.488
11	DW2	F	-29.470	64.941	67.440	41.523	8.343	0.990	0.940
12	DW2	C	-33.041	72.082	74.581	111.586	18.969	0.492	0.515
16	DW2	E, F, C	-28.914	67.827	71.993	54.703	10.146	0.836	0.870
18	GM	E	-36.126	78.253	80.752	187.004	5.793	0.042	0.000
19	GM	F	-31.213	68.427	70.927	120.121	27.414	0.689	0.607
20	GM	C	-35.543	77.086	79.585	219.149	4.335	0.066	0.072
24	GM	E, F, C	-28.404	66.808	70.974	36.949	32.989	0.917	0.954

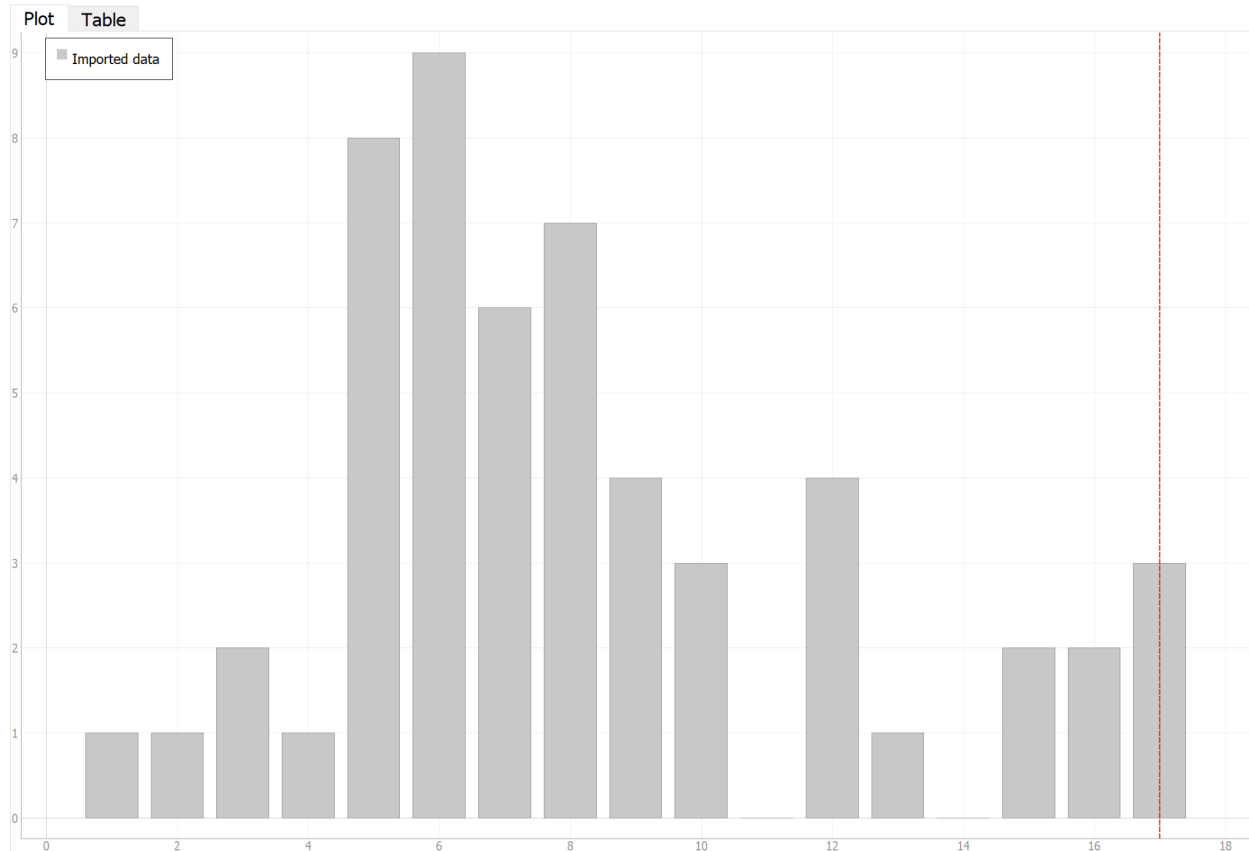
## Result of Range Analysis

*(An explanation of which part of the data is good for proceeding with the analysis)*

With the given failure rate and reliability of the data, data good to proceed within the analysis must fall within the line of a generally close estimate of the given failure data. Failure predictions that fall out of a general range of the original data can likely be considered an outlier and thus weighted lower than those with a closer estimate to the original. The reason for this is because the hazard functions are intended to model the given failure data and allow for the developer to view the tendency for the program to fail at any given point, whether that be at the time of the given failure data or a prediction in the future.

### Plots for Failure Rate and Reliability of the SUT





## Discussion on Decision Making for a given a Target Failure Rate

Based on the failure rate and reliability of the data given and predicted in the analysis, it would be optimal to aim for the target failure rate to be set at **approximately 5 failures every 3 intervals** going forward from the given data. This calculation was deduced based on the slopes of the given data compared with that of the predicted data viewed using the Geometric model, S Distribution model, and the Discrete Weibull model. Differences between these three hazard functions were taken into consideration based on their accuracy with the given failure data.

In the given data, failures generally followed failure rates tending to be around 3 failures every 4 seconds in the early intervals of the system (intervals 1-4), providing a failure rate of 0.75 failures per interval. However, in the later intervals, (intervals 5-10) this failure rate escalated to approximately 7 failures every 2 intervals, providing a failure rate of about 3.5 failures per interval. The cause of this escalation is likely due to the system being in its earlier development phases with overabundantly many unnoticed defects surfacing, causing this failure rate to become fairly bloated. This is made clear in the intervals following this (intervals 10-17) phase wherein the failures fall between 2 failures every interval, giving a failure rate of 2 failures per interval, and 3 failures every 2 seconds, giving a failure rate of 1.5 failures per interval.

The covariant defects also needed to be considered in the determination of the target failure rate as they each changed and affected the predicted failure rate's computation. Covariates represent defects in the system that can contribute in different quantities to the overall failure rate. For example, an S Distribution hazard function that takes into account only the covariate C for the interval of 11-13 has a failure rate of approximately 3 failures per interval. However, when considering covariates E, F, and C for the same S Distribution hazard function over that same interval predict a failure rate of about 1.5 failures per interval. In the end, the hazard function that has the failure rate with the closest prediction to the given data is chosen.

Finally, the target failure rate takes into account the future predictions for the intervals past the given data based on the prediction models. This data varied based on the model with the highest prediction being a failure rate of 2 failures per interval and the lowest prediction being close to 0.2 failures per interval. When considering the models that should be used for the prediction, more weight was placed on predictions that better matched the given failure data.

Given all of these considerations, a target failure rate of approximately 5 failures every 3 intervals was decided on. This rate provides a fair buffer for the given data's preceding failure rates but also expects a failure rate lower than its early failure rate spike, given that the failures should have leveled out over time, with an assumption of no new updates introduced to the data.

## Advantages and Disadvantages of Reliability Growth Analysis (Aashik)

This lab provided an experience to use the tools recommended/provided to learn about Reliability Growth Analysis. It also offered an extensive process with these tools to determine multiple pros and cons of the tools and their relationship with reliability growth analysis. Advantages of reliability growth analysis include collecting, modeling, analyzing, and interpreting data from data testing or from the field. In addition, reliability growth analysis allows us to analyze the growth process with different models provided by the tools.

Although the disadvantages of reliability growth analysis pale in comparison to its benefits, with a majority of them coming in the form of tools that are/were used. Specifically, the models that could be used were still only limited to the handful provided by the C-SFRAT tool (and was the only one that worked for this lab). Additionally, only certain types of metrics can be compared, so there is a limitation in that aspect. All in all, the advantages outweigh the disadvantages of reliability growth analysis, and the lab provided an extensive learning hands-on experience that increased our understanding of it.

[https://www.weibull.com/basics/growth.htm#:~:text=Reliability%20growth%20analysis%20is%20the,the%20field%20\(fielded%20systems\).](https://www.weibull.com/basics/growth.htm#:~:text=Reliability%20growth%20analysis%20is%20the,the%20field%20(fielded%20systems).)

## **RDC Testing**

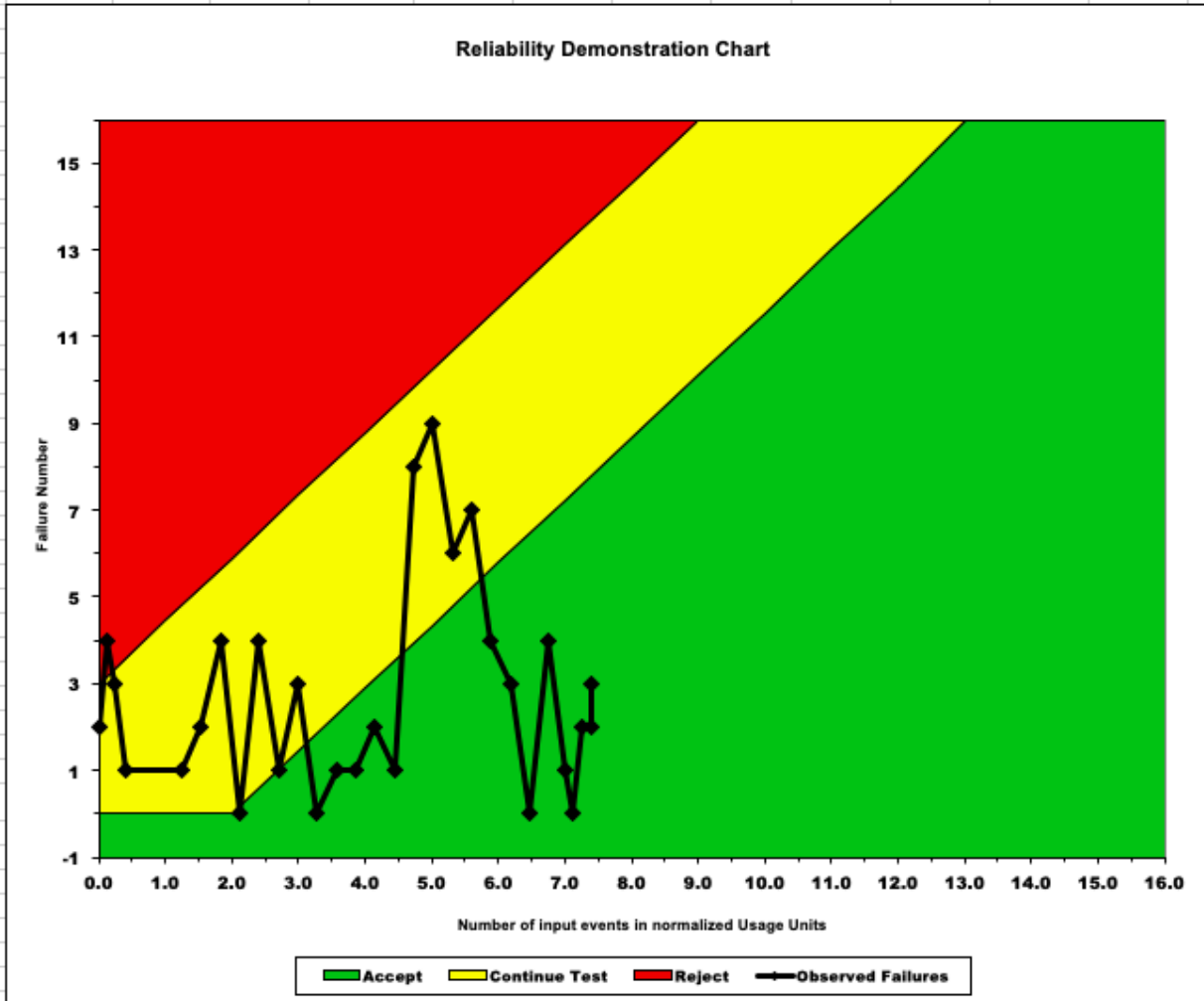
### **Three Plots**

The RDC chart shows whether or not the observed failures of the SUT have met the failure intensity object. The three regions on the chart indicate the different confidence levels of the SUT achieving the failure intensity objective. The red region indicates that the SUT cannot achieve the FIO, and thus must be rejected. The yellow region indicates that more tests needed to be conducted in order to determine if the SUT is acceptable or not. The green region indicates that the SUT will achieve the FIO. The 3 regions' boundaries are determined by taking into factors such as producer's risk threshold, customer's risk threshold, and the discrimination ratio. In the chart, the vertical axis represents the failure number, and the horizontal axis is normalized failure data (failure time divided by MTTF).

MTTF<sub>min</sub>

### Certification Testing Results for Project Example, System Demo 4.3

The FIO is 1 failures per 1000000 calls



Here the FIO is 1 failure per 1000000 calls.

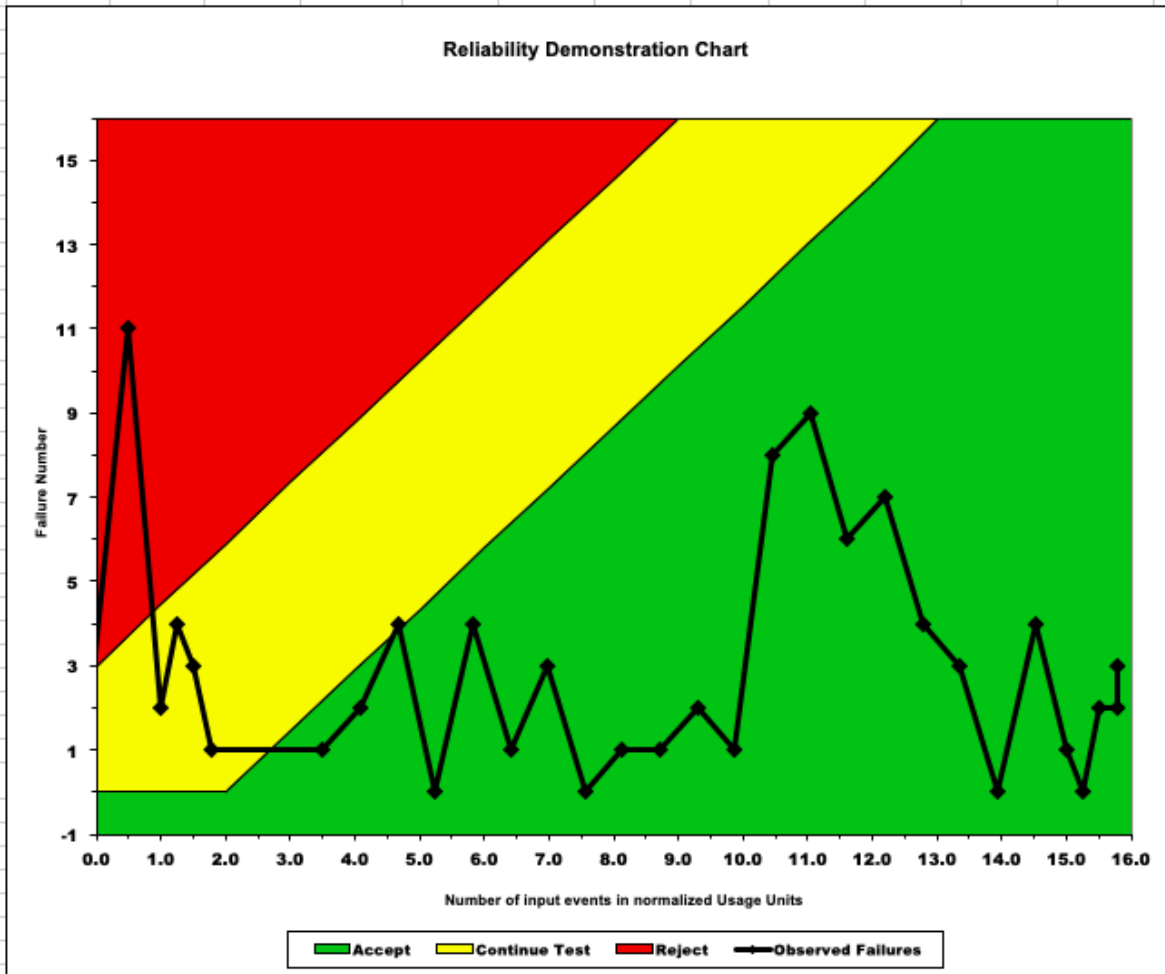
After analyzing the trend, the minimum MTTF<sub>min</sub> for which the SUT becomes acceptable is 6.89 because, after that value, all of the system's observed failures fall in the acceptable region.

The reason why the MTTF<sub>min</sub> is not 4.28 (the first time the SUT falls in the acceptable region) is that as more failures are observed, the SUT rises back into the continued region for testing. It is only after 6.89 that the SUT is solely in the acceptable area, which is why the MTTF<sub>min</sub> is 6.89.

Twice MTTFmin

### Certification Testing Results for Project Example, System Demo 4.3

The FIO is 2 failures per 1000000 calls



Here the FIO is 2 failures per 1000000 calls.

After analyzing the trend, the minimum MTTFmin for which the SUT becomes acceptable is 6.24 because, after that value, all of the system's observed failures fall in the acceptable region.

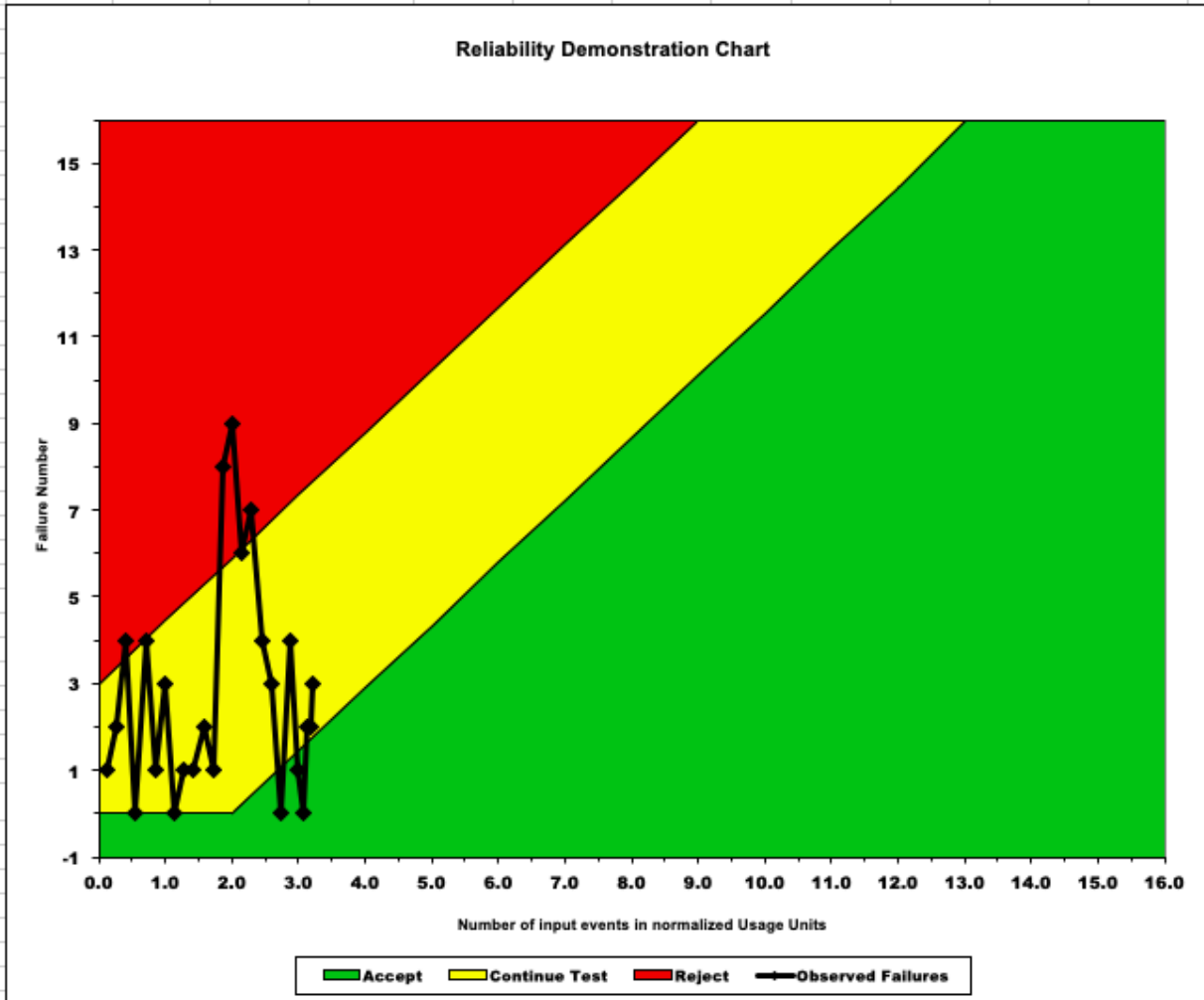
The reason why the MTTFmin is not 5.66 is because that point is on the boundary of the continuous and acceptable region. Since it is on the boundary, the next point, 6.24 is a better MTTFmin as everything after that is in the acceptable area.



Half MTTFmin

### Certification Testing Results for Project Example, System Demo 4.3

The FIO is 0.5 failures per 1000000 calls



Here the FIO is 0.5 failures per 1000000 calls.

After analyzing the trend, the MTTFmin cannot be concluded. More testing needs to be done on the SUT because the observed failures are in the continuous region. Interpreting the trend, it appears as if the MTTFmin for this failure data is quite large since a lot more testing is required.

No MTTFmin can be decided for this plot because even though the SUT becomes acceptable at points 3.74, 4.06, it rises back into the continue region (where more tests need to be conducted).

## Advantages and Disadvantages of RDC

A reliability demo chart (RDC) is a very cost-efficient and effective way of assessing system reliability. The assessment it produces can aid in making important decisions. Another advantage is that an RDC is a good visual way of looking at the system reliability metrics and is simple to interpret. However, a major limitation with RDC is that it cannot quantitatively assess the system reliability because it just shows the trend of changes that affect the reliability over time. Another disadvantage is that varying MTTF values for different scenarios are very tedious when it comes to the reliability demo chart.

## Discussion

### Comparison of the results with Part 1

The results in part 1 show higher rates of failure initially. The graph then starts to even out with failure rates not being as low averaging out at about 4. Furthermore, the results in part 2 consist of 3 plots. The first plot is with MTTF min and this plot shows the majority of points in the “continuous test” range as well as the “accept” test range. The second plot is twice MTTF min and in this plot, with initial failure numbers high (11) but the points further on being somewhat in the “continuous test” range and then many in the “accept” region. Lastly, the third plot in part 2 is the majority of the points in the yellow range and just a few in the green and red zones. This shows how the RDC changes with respect to the MTTF.

A discussion on the similarities and differences between the two techniques. Any lessons learned in this lab?

Some similarity between the two techniques is that the target failure rates can be visually interpreted in similar ways. This allows the graphs to represent similar information effectively. However, a difference is that the plot in part 1 can better quantify the data and the failure information whereas the RDCs are better visual representations. Some lessons we learned during this lab were mainly how to generate the graphs CSTRAT. Mainly how to interpret the best model for each type of failure distribution. In addition, we also learned how to make informed decisions about the software based on the chosen model to represent these failures.