Decisions Under Risk

– Quantitative Reliability in Action –

(Suitable for Self-Study)

Tim C. Adams, ASQ CRE

NASA Kennedy Space Center

Engineering Directorate

Technical Performance and Integration Division

Tim.Adams@NASA.gov

November 15, 2018 (rev a)

Presentation Objectives

- Provide real NASA examples (questions) where the program/project manager was faced with making a decision involving technical risk with uncertainty.
- Provide insights on how the Reliability-Risk Engineer responded to these questions (problems).
- Observe the uncertainty and the probabilistic method (as opposed to the traditional Physics-based deterministic) method in action.

Thinking Styles and Belief Systems

- **Big Q** vs. **Little q**
  - **Q**ualitative (expert opinion) vs. **q**uantitative (mathematical analysis)
  - Determinism vs. Probabilism

- "Basically, an **anti-empirical system** states that things may look like X, but in reality they are Y. Between us and reality is a screen of ideology."

  "God and Mankind: Comparative Religions," Robert Oden (former Professor and Chair, Department of Religion, Dartmouth College), The Great Courses, Lecture 2, 1998

- **Uncertainty** does not imply no knowledge, but it does imply the exact outcome is not completely predictable. Most observable phenomena contain a certain amount of uncertainty.

NASA Ideology: Safety and Mission Assurance

- "Don't bring me a perfect answer after launch."
  A NASA Johnson Space Center Manager's directive to his safety and mission assurance engineers
- Thus, answers (e.g., forecasts) can be probabilistic (not certain).
- In engineering assurance, there are four **types of uncertainty**.
  1. Aleatory Uncertainty
  2. Parameter (Statistical) Uncertainty
  3. Model Uncertainty
  4. Completeness Uncertainty
- **Aleatory uncertainty** deals with randomness and observed quantities (e.g., distance and time).
- The other three types are **epistemic uncertainty** that represents the state of knowledge and deals with non-observable quantities (e.g., failure rates and model assumptions).

Engineering Assurance (Specialty Engineering)*

Safety: Freedom from accident and loss	**Usability**: Human interfaces	**Supportability and Serviceability**: Service throughout the planned life cycle
Reliability: Likelihood of having an uptime (failure-free) state for a stated duration or load	**Maintainability**: Likelihood of returning to an uptime state during maintenance or repair	**Availability**: Likelihood a repairable item has an uptime state; f (**R**, **M**) = **A**
Producibility: Ease and economy of producing or manufacturing	**Affordability**: Total cost of ownership and not only system acquisition cost	**Disposability**: Disassembly and disposal (environmental stewardship)

***Engineering assurance** (as opposed to design engineering and engineering management) identifies and addresses issues and hazards early (i.e., during design, not during operation).

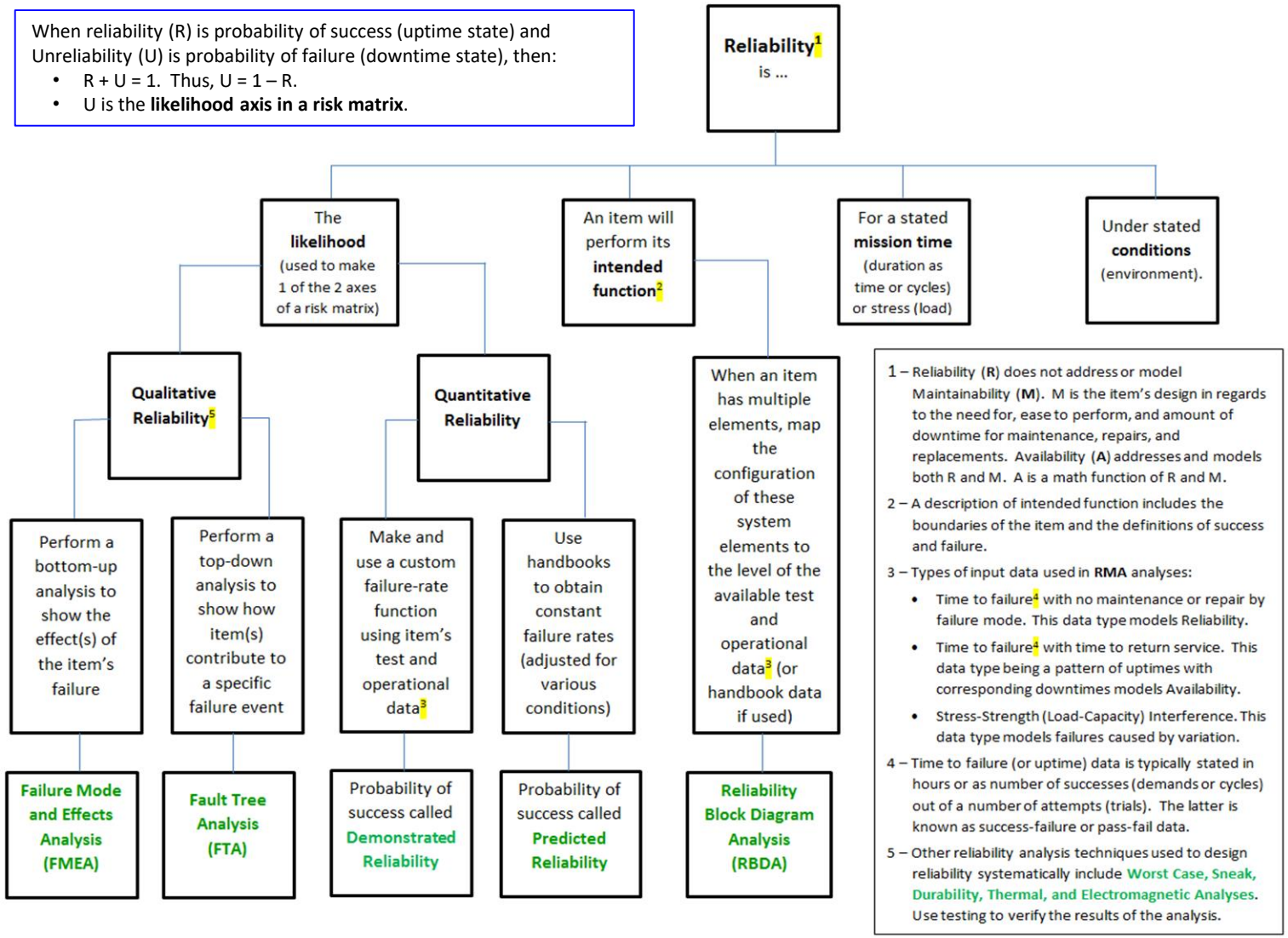
Reliability vs. Risk

- **Reliability** is the likelihood an item will perform its intended function for a specified period of time (or number of demands or load) with no failure under stated conditions.
- The measure for "not reliable" combined with the measure for "not safe" make a **risk** (potential loss) measure or point in a matrix.
- Understanding and prioritizing risk helps managers and engineers to make **"risk-informed" decisions**.

Reliability: From Concept to Products

When reliability (R) is probability of success (uptime state) and Unreliability (U) is probability of failure (downtime state), then:

- R + U = 1. Thus, U = 1 – R.
- U is the **likelihood axis in a risk matrix**.



1 – Reliability (**R**) does not address or model Maintainability (**M**). M is the item's design in regards to the need for, ease to perform, and amount of downtime for maintenance, repairs, and replacements. Availability (**A**) addresses and models both R and M. A is a math function of R and M.

2 – A description of intended function includes the boundaries of the item and the definitions of success and failure.

3 – Types of input data used in **RMA** analyses:

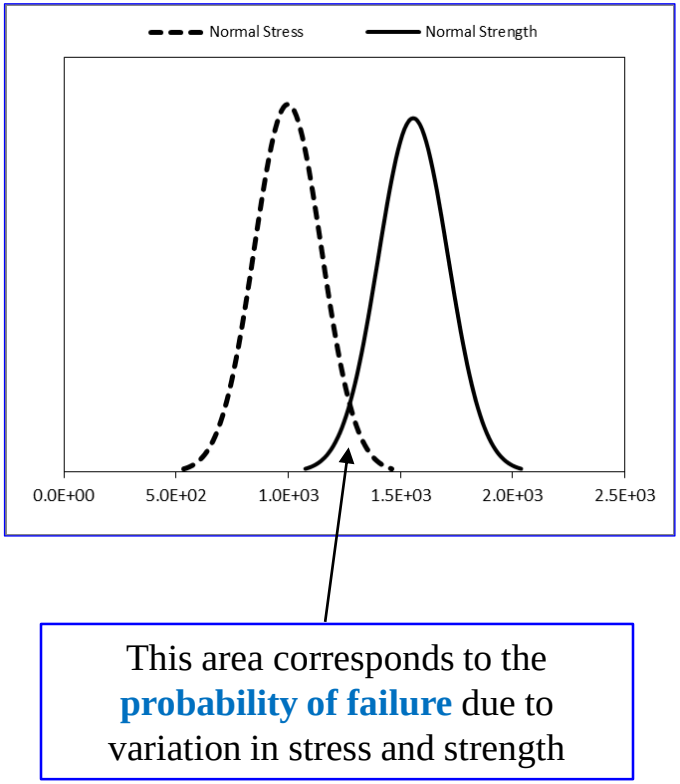
- Time to failure[4] with no maintenance or repair by failure mode. This data type models Reliability.
- Time to failure[4] with time to return service. This data type being a pattern of uptimes with corresponding downtimes models Availability.
- Stress-Strength (Load-Capacity) Interference. This data type models failures caused by variation.

4 – Time to failure (or uptime) data is typically stated in hours or as number of successes (demands or cycles) out of a number of attempts (trials). The latter is known as success-failure or pass-fail data.

5 – Other reliability analysis techniques used to design reliability systematically include **Worst Case, Sneak, Durability, Thermal, and Electromagnetic Analyses**. Use testing to verify the results of the analysis.

Data Types used in Reliability Analyses

- **Item under study** can be hardware, software, orgware (humans, human processes), or any combination.
- Quantitative reliability typically encounters **three types of data**:
  1. Time-based (clock) data
     - Continuous
     - Weibull (uptime), Log-normal (downtime).
  2. Event-based (demand) data
     - Discrete
     - Binomial (x or more), Poisson (x or less).
  3. Stress (load) and strength (capacity) data
     - For example, use normal distribution's mean and standard deviation of the item's stress and item's strength to calculate the **probability of failure**, the overlap in the distributions' tails.
     - Note: A safety factor does not characterize the item's uncertainty in stress and strength.



Risk

Risk as a Concept

- Risk is potential loss or potential gain.
- Thus, risk is the uncertain deviation (delta) in the execution of a management plan.
- Reference: ISO 31000.
- When limited to potential loss:
  - Risk is a qualitative or quantitative estimate of the potential loss occurring due to natural or human activities.

Risk as an Operation

- Both potential loss and gain:

  Actual results* = *Planned results* +/- *Risk

- Only potential loss:
  - **Scenario x**: What can go wrong?
  - **Likelihood for x**: What is the probability it will happen?
  - **Consequence for x**: What is the impact if it did happen?
  - **Risk measure for x** = (likelihood) * (Consequence). This assumes consequence is measurable.

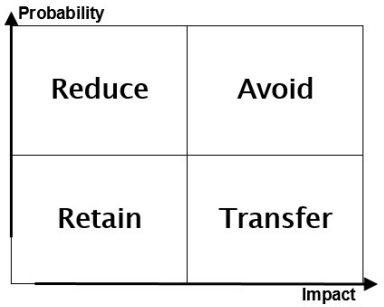
Good Decision = Good Information * Enough Processing Time

- "To a great extent, the successes or failures that a person experiences in life depend on the decisions that he or she makes.
- The person who managed the ill-fated the space shuttle Challenger mission is no longer working for NASA.
- The person who designed the top-selling Mustang became the president of Ford.
- Why and how did these people make their respective decisions?"

Quantitative Analysis for Management, 6th Ed, Render & Stair

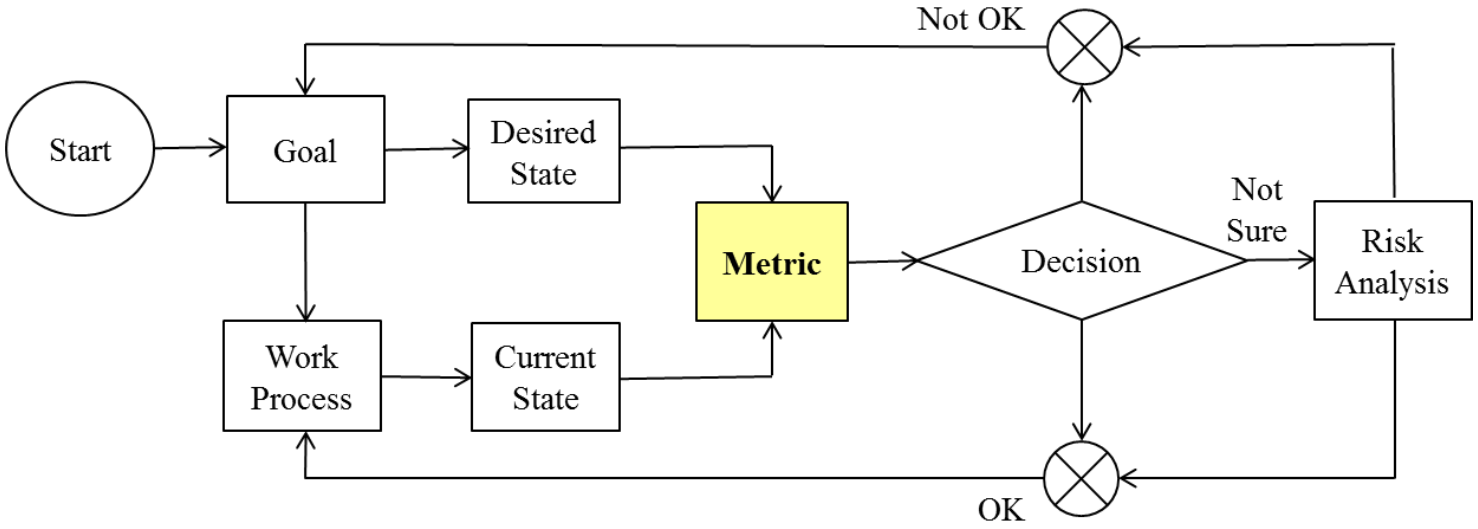
Five Ways: Deciding about Risk

1. **Accept** (retain, engage, fight).
2. **Avoid** (run, flight).
3. **Hold** (freeze; get information to later select one of the other four).
4. **Mitigate** (change something to reduce the risk; countermeasure).
5. **Transfer** (share the risk).



Systems Thinking and Measuring Performance

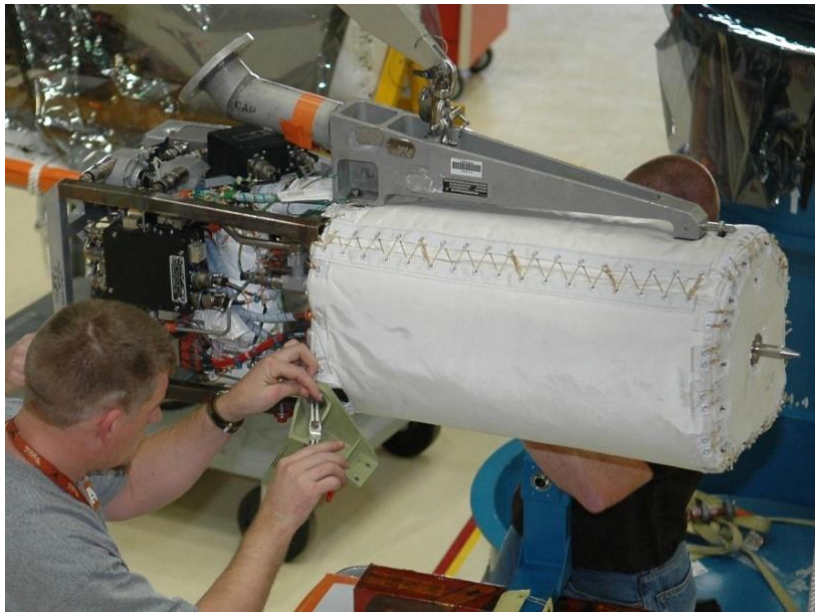
- **System**: A collection of different elements that together produce results not obtainable by the elements alone.
- **Metric**: The comparison of the current state to the desired state.
- **System Metrics**: **E**ffectiveness, **e**fficiency, and **a**ppeal (**E-e-a**).



Example #1 – The Problem

- A recent system failure caused major embarrassment as well as much expense. Should this system be replaced with new technology or upgraded?
- If upgraded, identify the system elements causing the trouble and the required reliability.

Example #1 – This problem occurred with a Space Shuttle Orbiter's Fuel Cell



A fuel cell on a Space Shuttle Orbiter caused a minimum duration flight (MDF) during STS-83. In addition to the MDF, a previous launch delay and numerous maintenance actions during "vehicle turnaround" made this system a serious candidate for improvement.

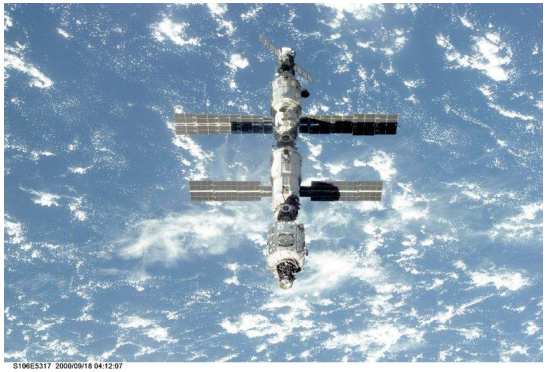
Example #1 – Reliability in Action

- A detailed reliability and maintenance (R&M) analysis and assessment report on all fuel cell line replaceable units (LRUs) from the STS-26 to STS-85 time period was completed.
- This R&M assessment was instrumental in the decision to change regulator material in all fuel cell LRUs for $12M instead of replacing with a new design estimated at $50M.

Example #2 – The Problem

- During a final review of a system prior to shipment, questionable test data appears on one of the system's components.
- Assuming all other system elements are believed to perform as expected, what is the risk of shipping as is?
- In particular, what is the likelihood the system will not work (i.e., perform to meet its minimum requirements)?

Example #2 – This problem occurred with the International Space Station's Gyroscopes as a payload

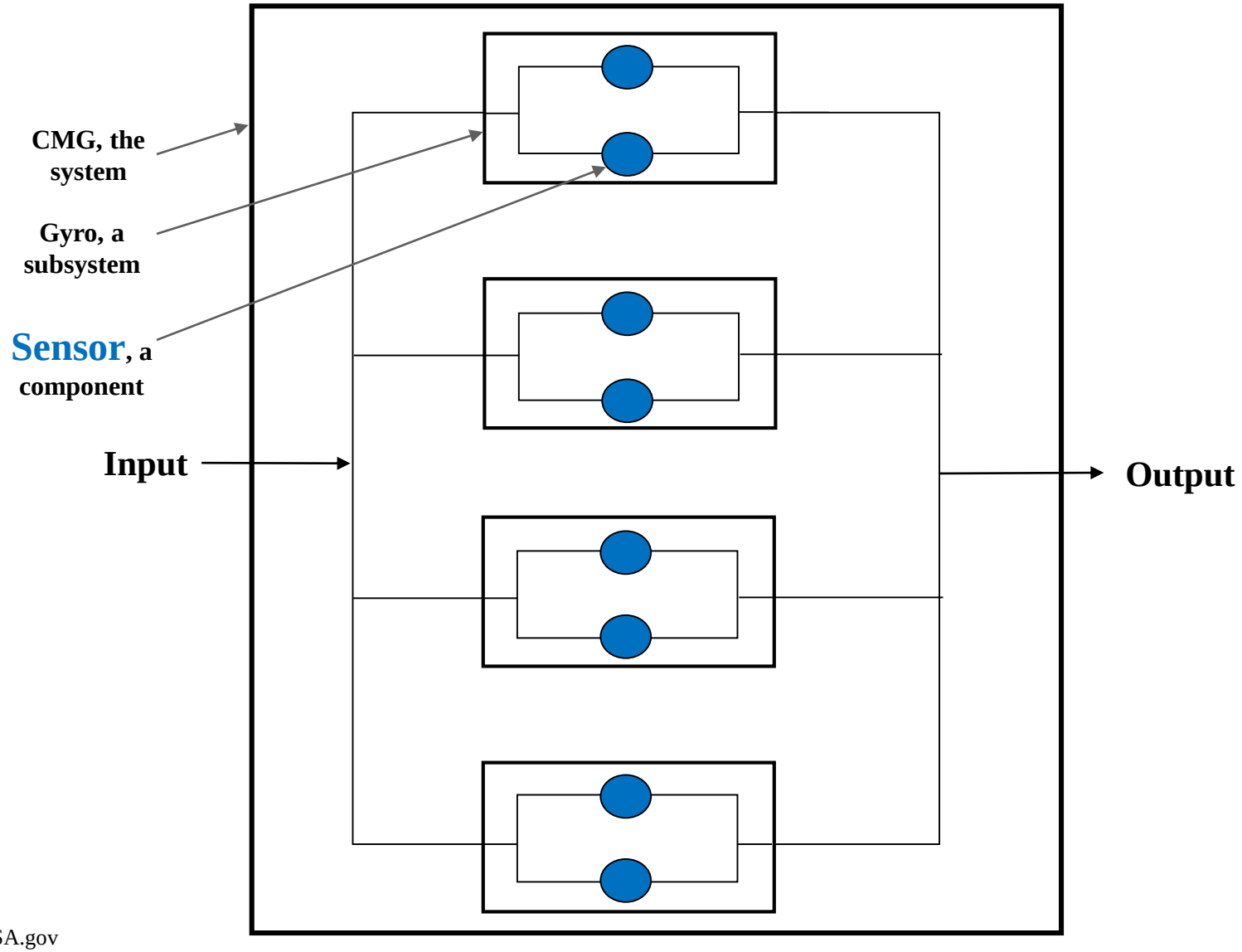


Sept 2000

Nov 2009

STS-118 Canadian Space Agency Astronaut and Mission Specialist Dafydd "Dave" Williams is attached to an Adjustable Portable Foot Restraint on the end of the Space Station Remote Manipulator System Canadarm 2 as it transports the new Control Moment Gyroscope (CMG) to the External Stowage Platform 2 for temporary stowage.

Example #2 – System Configuration for ISS Gyros



Example #2 – Reliability in Action

- Vendor test data generated a concern with the expected reliability of the Hall resolvers (sensors). These sensors are used on the Control Moment Gyros (CMGs), a Space Shuttle payload planned for Space Station's Flight 3A.
- A reliability analysis presented to the Space Station Control Board (SSCB) showed there was a 4-6% chance of not having a minimum CMG. A **minimum CMG** is at least 1 of 2 sensors working in each gyro and at least 2 of the 4 gyros working for 5, 7, or 9 thermal cycles.
- The uncertainty in the cycle count occurred because the sensor heaters would not be available on the Space Station until after the sensors experienced an estimated range (uncertain number) of thermal cycles.
- The consequence of not having a minimum CMG meant that 6 metric tons of propellant for a reboost would have to be consumed at a cost of $100 million.
- The decision was made to ship the CMGs as planned--which proved to satisfy the mission requirements both for the short and long terms.

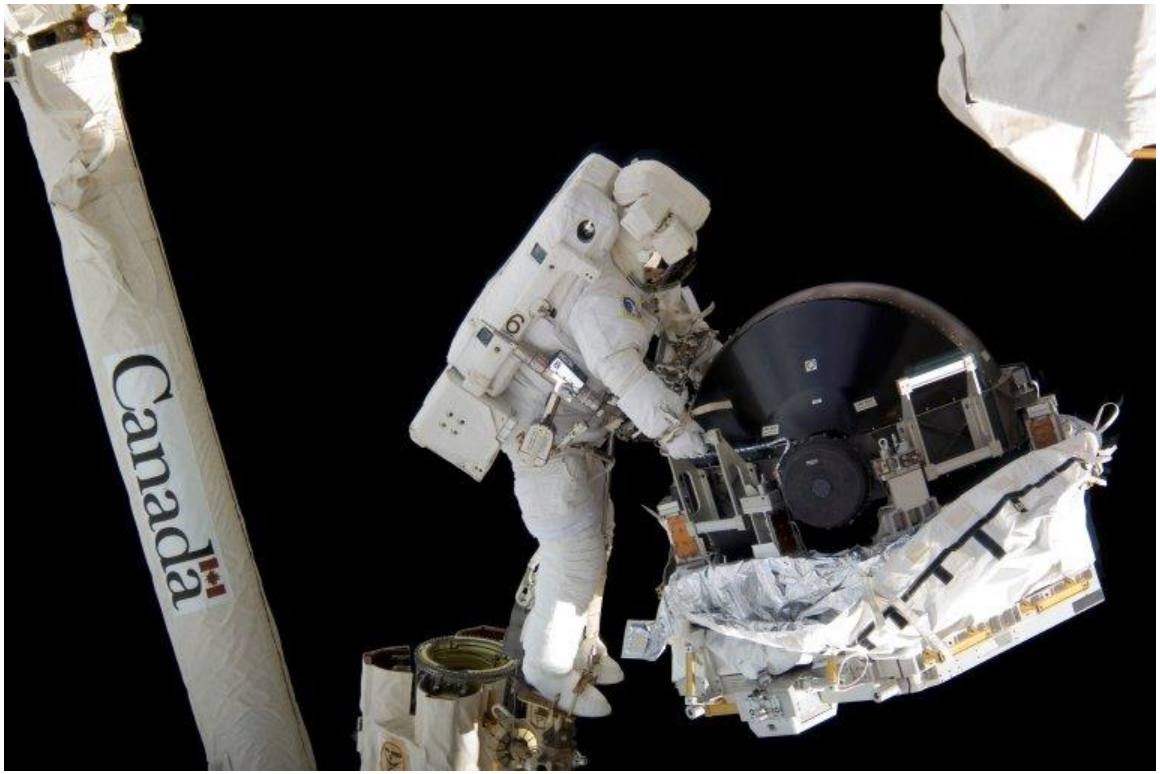
Example #3 – The Problem

- The test director wants to know, can testing stop after receiving no failures in 360 tests on a life-critical item?
- In particular, does this testing certify that the item is safe?

Example #3 – In particular, interpret test data

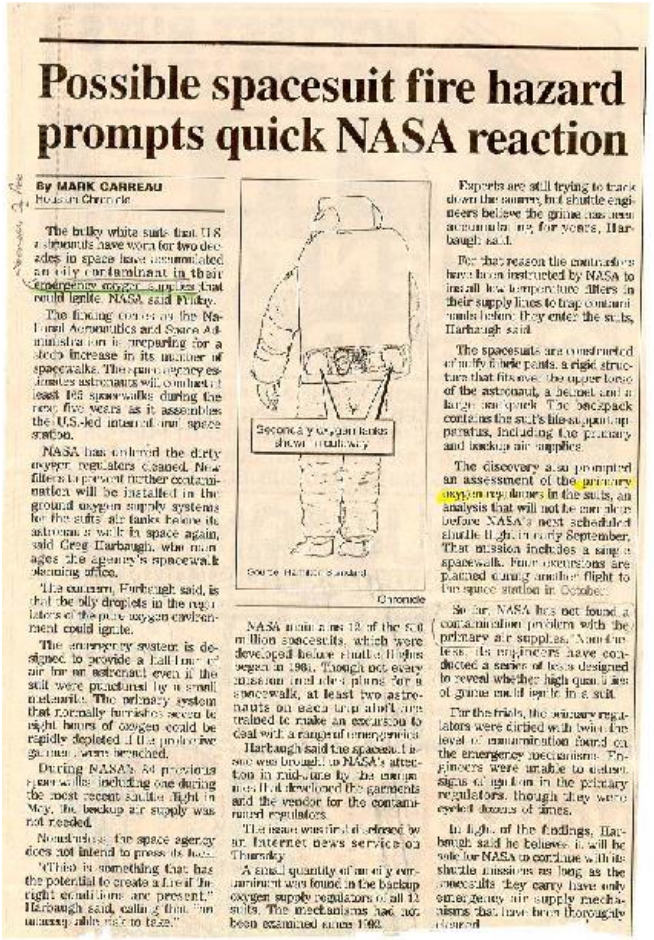
- The White Sands Test Facility (WSTF) conducted 360 tests to determine if ignition would occur during the presence of a small quantity of hydrocarbon oil in 100% oxygen under adiabatic compression, the compression heating of oxygen.
- None of the WSTF tests produced ignition. These tests were in response to a hydrocarbon-oil contaminate found in the Life Support System (PLSS) used in the Extravehicular Mobility Unit (EMU).

Example #3 – This problem pertained to the Astronaut's Extravehicular Mobility Unit



Extravehicular Mobility Unit (EMU) is an independent system that provides environmental protection, mobility, life support, and communications for a Space Shuttle or International Space Station (ISS) crew member to perform extra-vehicular activity (EVA) in earth orbit.

Example #3 – This problem made the news!



Possible spacesuit fire hazard prompts quick NASA reaction

By MARK CARREAU
Houston Chronicle

The bulky white suits that U.S. astronauts have worn for two decades in space have accumulated an oily contaminant in their emergency oxygen supplies that could ignite, NASA said Friday.

The finding comes as the National Aeronautics and Space Administration is preparing for a sharp increase in its number of spacewalks. The space agency estimates astronauts will conduct at least 168 spacewalks during the next five years as it assembles the U.S.-led international space station.

NASA has ordered the dirty oxygen regulators cleaned. New filters to prevent further contamination will be installed in the ground oxygen supply systems for the suits' six tanks before its astronauts walk in space again, said Greg Harbaugh, who manages the agency's spacewalk planning office.

The concern, Harbaugh said, is that the oily droplets in the regulators of the pure oxygen environment could ignite.

The emergency system is designed to provide a half hour of air for an astronaut even if the suit were punctured by a small meteorite. The primary system that normally furnishes seven to eight hours of oxygen could be rapidly depleted if the protective garment were breached.

During NASA's 84 previous spacewalks, including one during the most recent shuttle flight in May, the backup air supply was not needed.

Nonetheless, the space agency does not intend to press its luck.

"This is something that has the potential to create a hazard if the right conditions are present," Harbaugh said, calling that "an unacceptable risk to take."

Source: Hamilton Standard

Chronicle

NASA maintains 12 of the $11 million spacesuits, which were developed before shuttle flights began in 1981. Though not every mission includes plans for a spacewalk, at least two astronauts on each trip aboard are trained to make an excursion to deal with a range of emergencies.

Harbaugh said the spacesuit issue was brought to NASA's attention in mid-June by the contractor that developed the garments and the vendor for the contaminated regulators.

The issue was first reported by Aviation Internet news service on Thursday.

A small quantity of oily contaminant was found in the backup oxygen supply regulators of all 12 suits. The mechanisms had not been examined since 1992.

Experts are still trying to track down the source, but shuttle engineers believe the grime has been accumulating for years, Harbaugh said.

For that reason the contractors have been instructed by NASA to install low-temperature filters in their supply lines to trap contaminants before they enter the suits, Harbaugh said.

The spacesuits are constructed of nylon fabric pants, a rigid structure that fits over the upper torso of the astronaut, a helmet and a large backpack. The backpack contains the suit's life-support apparatus, including the primary and backup air supplies.

The discovery also prompted an assessment of the primary oxygen regulators in the suits, an analysis that will not be complete before NASA's next scheduled shuttle flight in early September. That mission includes a single spacewalk. Four excursions are planned during another flight to the space station in October.

So far, NASA has not found a contamination problem with the primary air supplies. Nonetheless, its engineers have conducted a series of tests designed to reveal whether high quantities of grime could ignite in a suit.

For the trials, the primary regulators were ditched with twice the level of contamination found on the emergency mechanisms. Engineers were unable to detect signs of ignition in the primary regulators, though they were cycled dozens of times.

In light of the findings, Harbaugh said he believes it will be safe for NASA to continue with its shuttle missions as long as the spacesuits they carry have only emergency air supply mechanisms that have been thoroughly cleaned.

Example #3 – Reliability in Action but limited

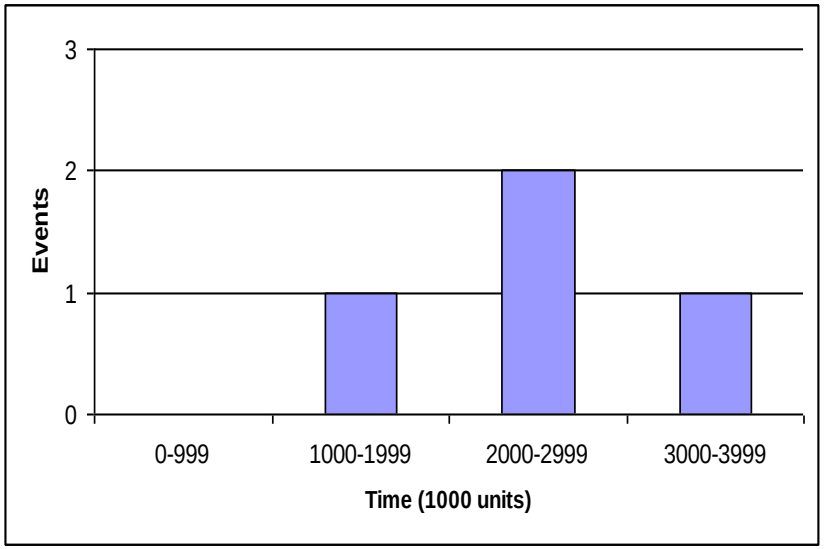
- Method 1 used **Classical Test Statistics** to determine the maximum failure probability with a high degree of statistical confidence. This failure probability did not meet the program's failure-probability goal. Thus, *more testing* would have been required if only this analysis method ("little q") was used for decision making.
- Method 2 used **Ancillary Data** (i.e., similar test data) to identify a boundary for ignition and no ignition. This method did not address heat loss and was *not sufficient for decision making*.

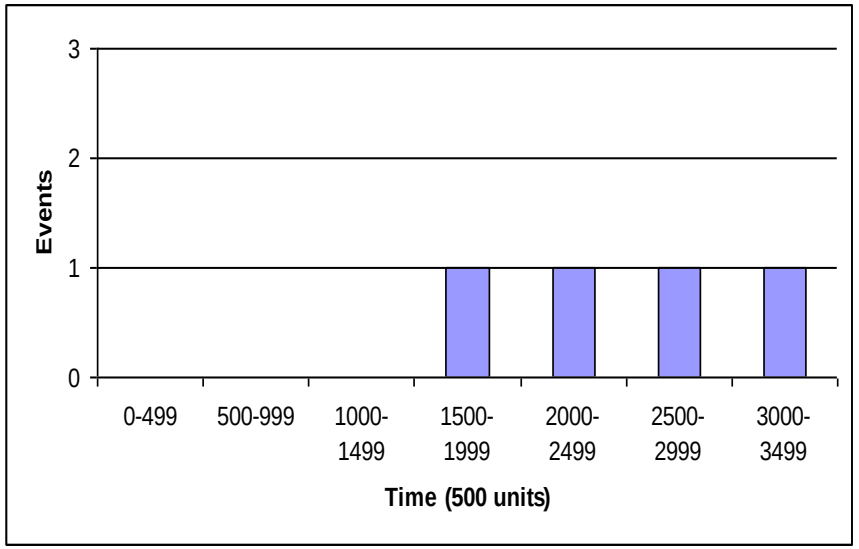
Example #3 – Reliability in Action but limited

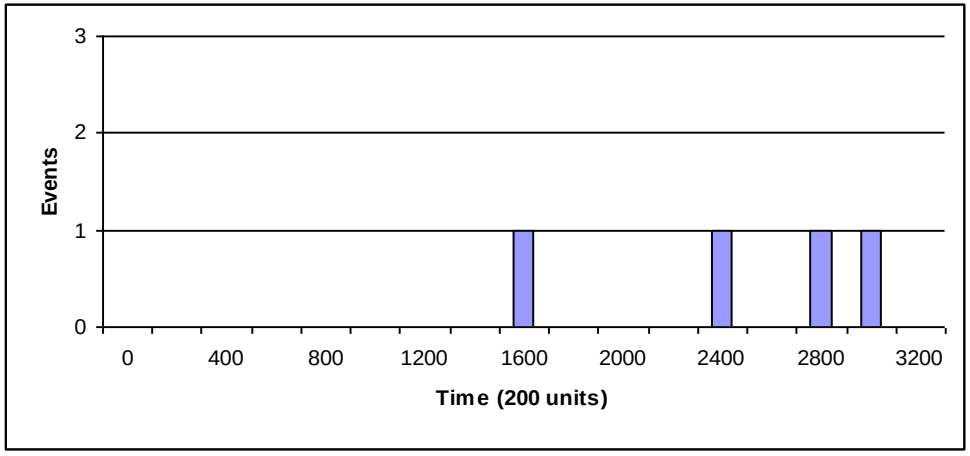
- Method 3 used the **Arrhenius Reaction Rate Model**. This model (pseudo-"big Q") adjusted the failure probability found in the first method since all WSTF testing was done under stressed conditions (higher pressure). *The failure-probability goal was surpassed under certain assumptions*.
- Method 4 used **Combustion Physics** (Semenov equations) to address the heat loss not addressed in Method 2. It was found that the reaction rate was not fast enough to cause ignition in the PLSS. Thus qualitatively ("big Q"), the *failure-probability goal was believed (decided) to be satisfied*.

Example #4 – What is the trend for these discrete events?

That is, for each chart, is the trend increasing, decreasing, or constant?







Example #4 – Determine the trend without graphing

- All graphs on the previous page use the same data!
- To test for a trend without graphing, use the **Laplace Test** (**U**), a test statistic.
- As a discrete event, 1600, the first event in the third histogram, could mean:
  - From a problem reporting database, the number of days from 01/01/09 (the start date of the trend period under study) to 05/20/13 (date of first event for item xyz).
  - From a machine-hour meter, the number of hours a machine (e.g., server) has operated from the time it was new and restarted after each upgrade or refurbishment.
- The Laplace Test applied to the four absolute times (1600, 2400, 2800, and 3000) and 3800 as the selected length of the observed period, make **U**, the test statistic, is +1.00 (< 0 is decreasing trend, 0 is no trend, >0 is increasing trend).
- Formally: Since **U** = +1.00, then z = 1.00 in a normal distribution table indicates the area to the left of z is 0.8413. When a **statistical hypothesis** is based on a one-tailed test and it is decided to reject the null hypothesis (H_o: data has no trend) with ~84% confidence, the research hypothesis (H_a: data has trend, up since z > 0) is assumed true. The probability this decision is the wrong (Type I error) is ~16%. This type of error is denoted α and is called producer's risk.

Example #4 – This "dashboard" reports 182 trends as a score for 26 systems

ACTIVITY, TREND, & RISK RESULTS FOR PROBLEM RECORDS FROM XXXXXX PRACA DURING 07-01-95 TO 06-30-00

MAJOR SYSTEMS FOR XXXXXX		Data Set 6: All Problem Records. The period is from 07-01-95 to 06-30-00.		Data Set 5: Invalid Problems. An invalid problem is a duplicate call-in or a "non-reportable" problem.		Data Set 4: Problems Pending Final Disposition. A pending problem is an open problem or a problem on a tracking list.		Data Set 3: Problems Resulting in No Corrective Action with or without Remedial Action. Remedial action is a repair or fix to one LRU. Corrective action identifies root cause and actions to prevent recurrence on all LRUs.			Data Set 2: Problems Resulting in Recommended Corrective Action that Occurred During Non-Critical Periods. Non-critical periods are initial ATP, qual, laboratory, manufacturing, bench-test maintenance, and development.			Data Set 1: Problems Resulting in Recommended Corrective Action that Occurred During Critical Periods. Critical periods are non-initial ATP, OMRSD, and flight.			Data Sets 1-3 Combined: Dispositioned Valid Problems. This data set, consisting of only valid and closed problems, provides a historical basis for the three measures below.				Data Sets 1-4 Combined: All Valid Problems. This data set, consisting of only valid problems with some still pending, is used to measure risk via a function of problem count, trend, and criticality.					
Major System	All Problem Records in xxx PRACA from 01-01-81 to 06-30-00	Count (Percent of Column Total Count)	Trend Score†	Count (Percent of Data Set 6 System Count)	Trend Score†	Count (Percent of Data Sets 1-4 System Count)	Average Age (months)	Count (Percent of Data Sets 1-4 System Count)	Trend Score†	Failure Criticality Measure[2,3]	Count (Percent of Data Sets 1-4 System Count)	Trend Score†	Failure Criticality Measure[2,3]	Count (Percent of Data Sets 1-4 System Count)	Trend Score†	Failure Criticality Measure[2,3]	Count (Percent of Column Total Count)	Unexplained Anomalies: Count (Percent of Data Sets 1-3 System Count)	Unexplained Anomalies: Trend Score†	Count (Percent of Column Total Count)	Trend Score†	Failure Criticality Measure[2,3]	Risk Measure as a Percent of Total X	Chart Page Number	Narrative Page Number
AXX	1652	319 (9.4%)	-4.73	31 (9.7%)	0.12	71 (24.7%)	24.6	122 (42.4%)	-4.64	0.54	56 (19.4%)	-5.25	0.45	39 (13.5%)	-2.68	0.52	217 (9.3%)	35 (16.1%)	-2.49	288 (9.4%)	-5.02	0.52	23.8%		3-32
BXX	2159	337 (9.9%)	1.21	37 (11.0%)	-0.97	81 (27.0%)	14.6	143 (47.7%)	-3.25	0.05	50 (16.7%)	1.21	0.06	26 (8.7%)	-2.67	0.06	219 (9.3%)	26 (11.9%)	-2.12	300 (9.8%)	1.62	0.05	15.9%		3-34
CXX †	2721	218 (6.4%)	1.35	29 (13.3%)	1.46	36 (19.0%)	15.2	142 (75.1%)	-0.85	0.14	2 (1.1%)		0.00	9 (4.8%)	-2.24	0.10	153 (6.5%)	12 (7.8%)	-1.42	189 (6.2%)	0.88	0.15	13.3%		3-34
DXX	228	228 (6.7%)	1.24	27 (11.8%)	0.39	57 (28.4%)	13.9	76 (37.8%)	-1.42	0.04	58 (28.9%)	-3.05	0.05	10 (5.0%)	-0.24	0.06	144 (6.1%)	0 (0.0%)		201 (6.6%)	1.17	0.07	8.2%		3-35
EXX †	519	347 (10.2%)	-8.55	40 (11.5%)	-1.08	78 (25.4%)	25.1	208 (67.8%)	-11.04	0.10	9 (2.9%)	-2.74	0.10	12 (3.9%)	-2.11	0.08	229 (9.8%)	7 (3.1%)	-0.88	307 (10.0%)	-8.71	0.15	7.3%		3-32
FXX	251	52 (1.5%)	2.79	4 (7.7%)	1.11	22 (45.8%)	12.3	7 (14.6%)	-1.03	0.30	10 (20.8%)	0.55	0.62	9 (18.8%)	-1.19	0.56	25 (1.1%)	0 (0.0%)		48 (1.6%)	2.59	0.43	6.7%	3-24	3-30
GXX	1541	284 (8.3%)	-4.73	26 (9.2%)	-2.21	22 (8.5%)	15.6	166 (64.3%)	-4.19	0.08	43 (16.7%)	-1.51	0.06	27 (10.5%)	-4.39	0.08	236 (10.1%)	22 (9.3%)	-3.89	258 (8.4%)	-4.26	0.08	3.2%		3-35
HXX	635	96 (2.8%)	-0.01	17 (17.7%)	0.12	20 (25.3%)	10.4	25 (31.6%)	-0.85	0.18	15 (19.0%)	0.76	0.47	19 (24.1%)	-5.01	0.39	59 (2.5%)	2 (3.4%)		79 (2.6%)	-0.06	0.31	2.9%	3-25	3-29
IXX	599	130 (3.8%)	-3.90	8 (6.2%)	-0.30	29 (23.8%)	20.7	39 (32.0%)	-4.13	0.16	21 (17.2%)	-2.73	0.69	33 (27.0%)	-3.64	0.43	93 (4.0%)	10 (10.8%)	-4.18	122 (4.0%)	-3.95	0.33	2.9%		3-33
JXX	698	198 (5.8%)	-4.97	3 (1.5%)		55 (28.2%)	23.2	65 (33.3%)	-5.15	0.11	10 (5.1%)	-1.43	0.05	65 (33.3%)	-5.40	0.13	140 (6.0%)	34 (24.3%)	-5.97	195 (6.4%)	-4.86	0.12	2.5%		3-34
KXX	857	163 (4.8%)	-0.58	16 (9.8%)	-0.67	41 (27.9%)	21.7	58 (39.5%)	-0.92	0.07	23 (15.6%)	-2.46	0.08	25 (17.0%)	-1.35	0.14	106 (4.5%)	8 (7.5%)	-2.77	147 (4.8%)	-0.49	0.09	2.4%		
LXX	671	112 (3.3%)	0.08	16 (14.3%)	0.15	24 (25.0%)	16.9	34 (35.4%)	0.23	0.11	12 (12.5%)	-1.74	0.08	26 (27.1%)	-2.60	0.25	72 (3.1%)	4 (5.6%)	-2.19	96 (3.1%)	0.02	0.15	2.1%		
MXX	210	91 (2.7%)	1.15	7 (7.7%)	0.20	25 (29.8%)	18.0	39 (46.4%)	-1.21	0.09	15 (17.9%)	0.44	0.10	5 (6.0%)	-1.63	0.08	59 (2.5%)	0 (0.0%)		84 (2.7%)	0.85	0.10	1.6%		
NXX	97	19 (0.6%)	3.80	2 (10.5%)		6 (35.3%)	14.2	7 (41.2%)	3.38	0.29	2 (11.8%)		0.50	2 (11.8%)		0.50	11 (0.5%)	0 (0.0%)		17 (0.6%)	3.57	0.35	1.4%	3-26	3-31
OXX	1620	214 (6.3%)	-2.16	20 (9.3%)	-1.26	51 (26.3%)	20.1	121 (62.4%)	-3.94	0.02	7 (3.6%)	-1.61	0.00	15 (7.7%)	-1.93	0.01	143 (6.1%)	19 (13.3%)	-2.60	194 (6.3%)	-1.86	0.04	1.4%		
PXX	539	5 (1.9%)	-1.79	5 (7.7%)	0.65	4 (23.3%)	26.3	23 (38.3%)	-3.04	0.39	8 (13.3%)	-0.22	0.63	15 (25.0%)	-0.97	0.41	45 (2.0%)	0 (0.0%)		60 (2.0%)	-2.05	0.45	1.3%		3-31
QXX †	538	249 (7.3%)	-2.51	35 (14.1%)	-0.18	22 (10.3%)	16.2	150 (70.1%)	-4.39	0.04	13 (5.1%)	0.18	0.01	29 (13.6%)	-0.55	0.04	192 (8.2%)	12 (6.3%)	-0.31	214 (7.0%)	-2.63	0.04	1.2%		
RXX †	547	76 (2.2%)	-0.94	7 (9.2%)	1.16	25 (36.2%)	18.0	35 (50.7%)	-4.06	0.15	3 (4.3%)		0.10	5 (8.7%)	-0.14	0.23	44 (1.9%)	0 (0.0%)		59 (2.3%)	-1.35	0.21	0.9%		3-30
SXX	244	35 (1.0%)	-0.34	3 (8.6%)		10 (31.3%)	14.4	16 (50.0%)	-1.07	0.47	4 (12.5%)	-2.54	0.77	2 (6.3%)		0.50	22 (0.9%)	3 (13.6%)		32 (1.0%)	-0.44	0.43	0.6%		3-33
TXX	690	59 (1.7%)	-0.53	6 (10.2%)	-0.51	12 (22.6%)	19.2	29 (54.7%)	-1.91	0.05	6 (11.3%)	0.16	0.07	6 (11.3%)	-0.18	0.07	41 (1.7%)	0 (0.0%)		53 (1.7%)	-0.39	0.08	0.3%		
UXX	811	4 (1.9%)	-1.79	5 (7.8%)	0.00	10 (16.9%)	20.9	41 (69.5%)	-3.53	0.05	4 (6.8%)	1.81	0.05	4 (6.8%)	-0.29	0.03	49 (2.1%)	4 (8.2%)	-0.93	59 (1.9%)	-1.86	0.05	0.2%	3-27	3-33
VXX	175	31 (0.9%)	-1.28	1 (3.2%)		3 (10.0%)	12.6	24 (80.0%)	-2.19	0.00	1 (3.3%)		0.00	2 (6.7%)		1.00	27 (1.2%)	0 (0.0%)		30 (1.0%)	-1.54	0.07	0.1%		
WXX	100	5 (0.1%)	-2.33	0 (0.0%)		0 (0.0%)		2 (40.0%)		0.50	0 (0.0%)			3 (60.0%)		0.33	5 (0.2%)	0 (0.0%)		5 (0.2%)	-2.33	0.40	0.0%		
XXX	121	15 (0.4%)	-1.17	5 (33.3%)	1.68	0 (0.0%)		8 (80.0%)	-2.19	0.09	1 (10.0%)		0.00	1 (10.0%)		0.10	10 (0.4%)	0 (0.0%)		10 (0.3%)	-2.62	0.08	0.0%		3-29
OTHER	15	2 (0.1%)		0 (0.0%)		2 (100.0%)	39.2	0 (0.0%)			0 (0.0%)			0 (0.0%)			0 (0.0%)	0 (0.0%)		2 (0.1%)		0.00			
TOTAL	20338	3409 (100.0%)	-8.04	350 (10.3%)	-0.48	716 (23.4%)	19.1	1580 (51.7%)	-14.98	0.13	373 (12.2%)	-5.93	0.21	390 (12.7%)	-10.25	0.22	2343 (100.0%)	198 (8.5%)	-9.15	3059 (100.0%)	-8.32	0.16	100.0%		

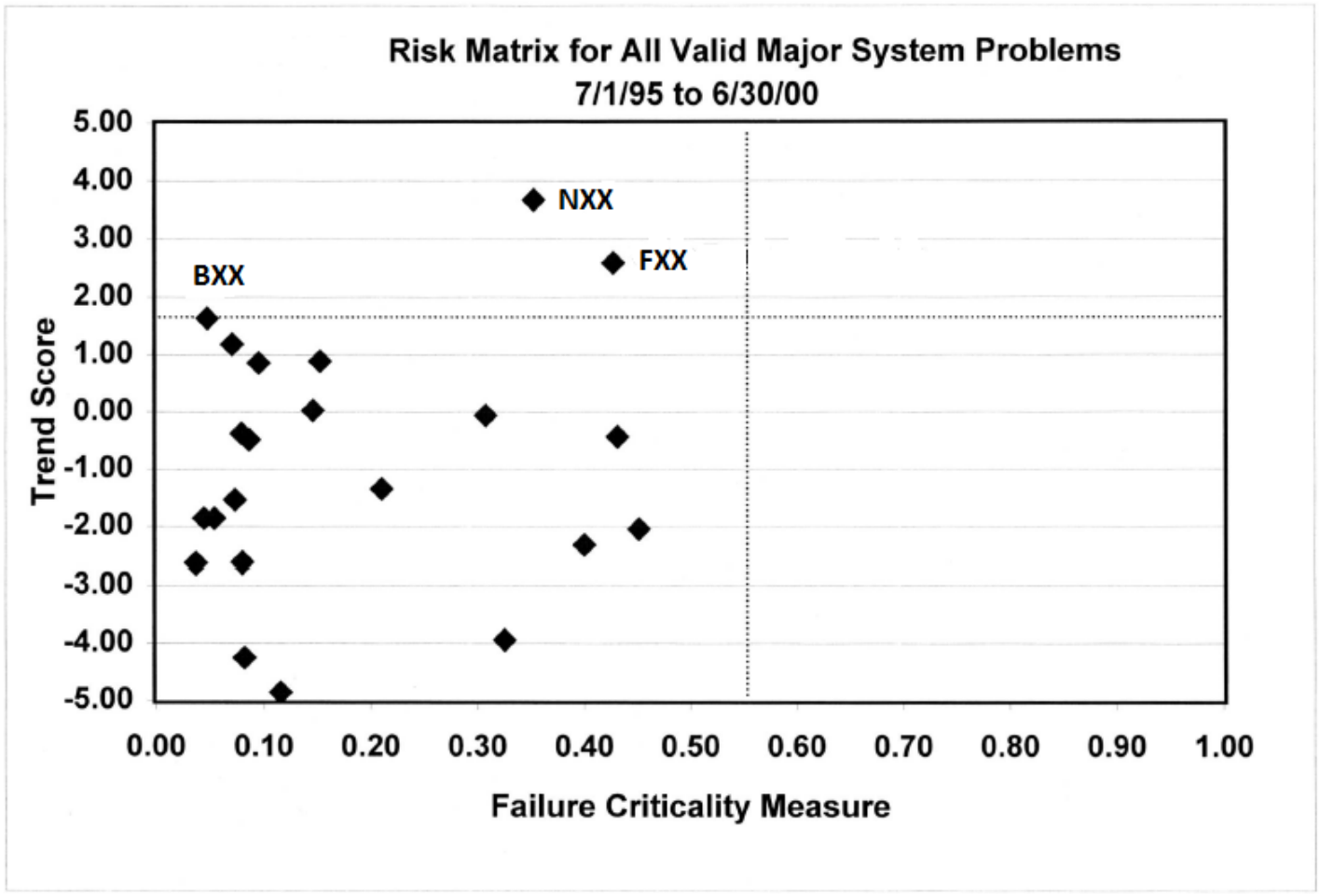
Notes: † An MOU(s) resulting in problems not being reported has been written on components in this system. See Appendix G for information on the impact of the MOU on problem activity and trends.

1. A trend score is calculated when the problem count is four or greater. 1.65 is the level of 90% confidence. A positive trend score is bad (increasing), while a negative trend score is good (decreasing).
2. The failure criticality measure is a density based on a problem's assigned failure criticality. All Crit 1/1 = 1.00, all Crit 1R = 0.10, and all other Crit's = 0.0. A failure criticality measure of 1.00 implies a higher risk than a failure criticality measure of 0.00.
3. The failure criticality measure for Data Sets 1-4 Combined represents the minimum failure criticality measure. As pending problems receive their failure criticality code, this measure could increase.

Trigger Finding Criteria	Trend Score	UA Trend Score (Data Sets 1-3 only)	Failure Criticality Measure	Average Age of Open Pending Cars	Problem Count Percentage				Risk Measure
Red (Most Undesired)	≥ 1.65	> 0.00	> 0.55, if trend score exists	> 24 Months	Invalid problems > 2 times the Total XXXXXX %	Pending problems > Total XXXXXX % + 10%	Problems resulting in no corrective action with or without remedial action > 80% and Count > 5	UA's > 2 times the Total XXXXXX %	Top 80% Contributors
Yellow (Undesired)	0.0 to 1.65	-1.00 to 0.00	0.32 to 0.55, if trend score exists	12 to 24 Months					

Example #4 – Risk matrix for previous report ("risk measure" not used)

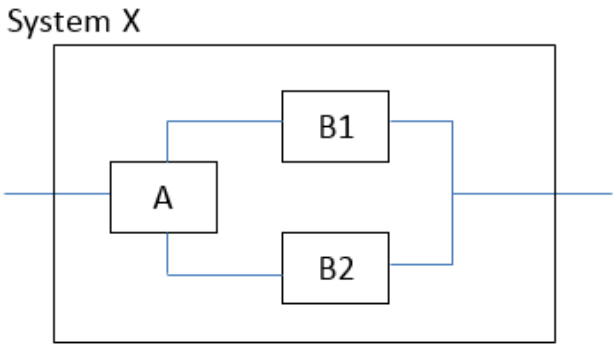
XXXXXX REPORT PRELIMINARY FINDINGS



Example #5 – System Reliability

1 - Given

a. System Description: System X must satisfy requirements A and B. B is dependent on the success of A. B has two independent elements, denoted B1 and B2. At least one of the two B elements must be successful. The diagram that follows is the block diagram for System X's two requirements.



Event	Reliability
A	90%
B1	80%
B2	70%

b. Data (*3 types*): (1) The table provides the point estimate values (not the probability distributions) for the likelihood of success (reliability) for each *requirement* based on experience. (2) If needed, assume the likelihood of the *initiating event* that triggers the need for System X equals one. (3) The quantity or measure for failure *consequence* at this time will be a placeholder since it is unknown.

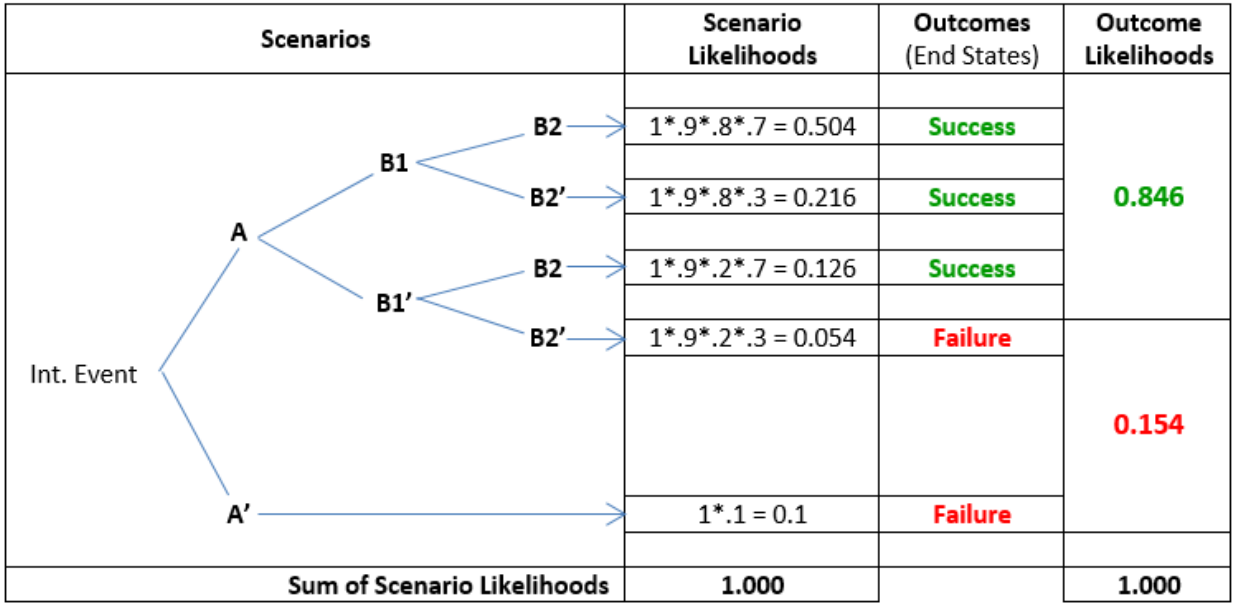
c. Problem Statement: What is the likelihood System X will successfully perform its intended function?

Example #5 – System Reliability

2 - Analysis

d. Outcome Statement: If S denotes success and S' denotes the complement of success (failure), and there is no degraded mission state (partial success), then S = A and (B1 or B2 given A).

e. Event Tree Method: Events A, B1, and B2 generate eight (2^3) possible scenarios. Scenarios need to be assessed for applicability. As given, since B is dependent on A, then $P(B_i|A') = P(A'|B_i) = 0$. Thus, P(A' and B_i) = 0. Therefore, events A' and B_i as are mutually exclusive (i.e., if A' occurs, then B_i cannot, and conversely). Also, since it is given events B1 and B2 are independent, then P(B1 and B2) = P(B1)*P(B2).

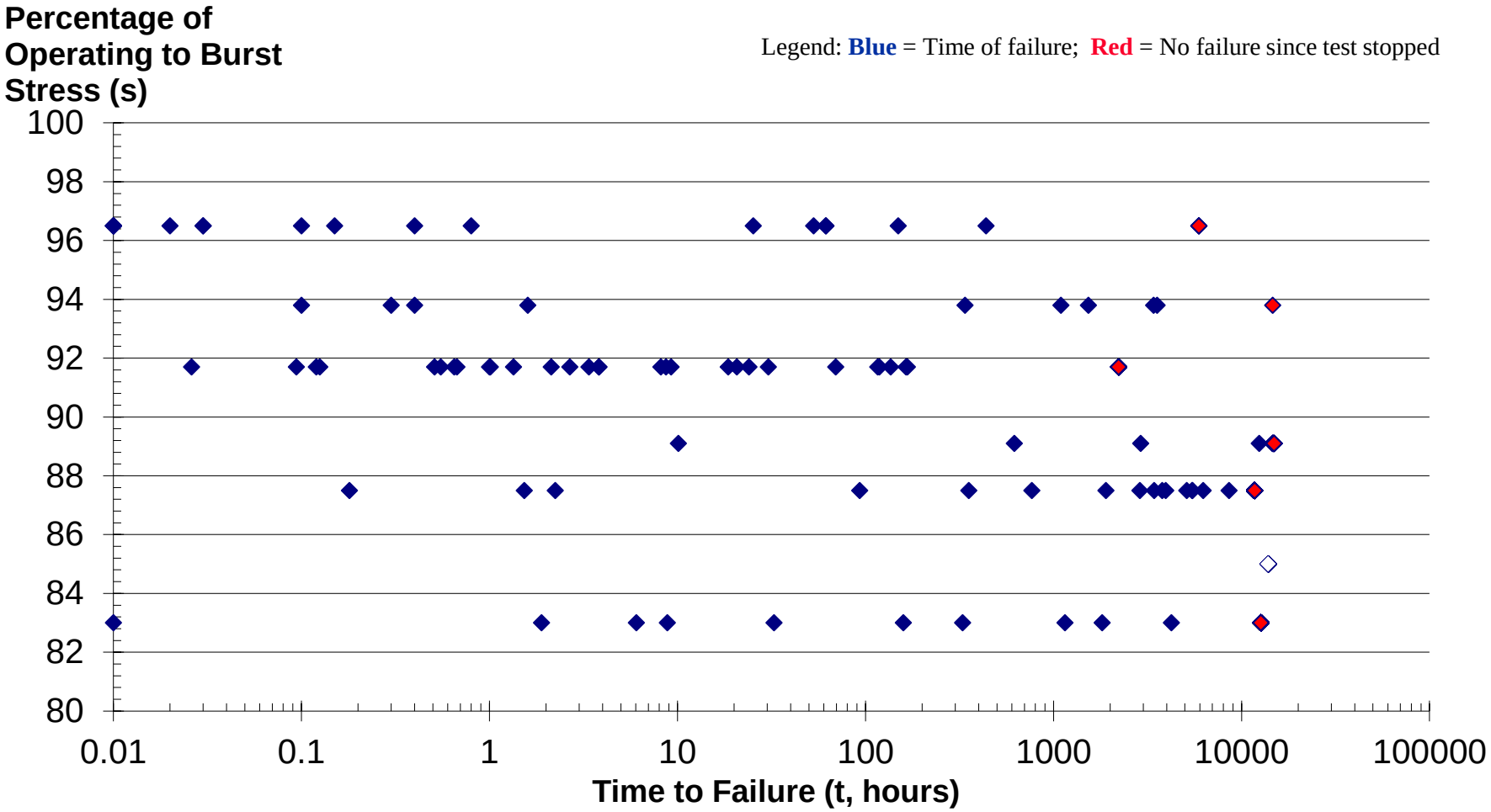


f. Formula Method:

P(System X Success) = P(A and [B1 or B2]) = P(A)*[1-(1-P(B1)*(1-P(B2)] =(0.9)*[1-(1-0.8)*(1-0.7)] = (0.9)*[1-(0.2)*(0.3)] = (0.9)*[0.94] = **0.846**.

P(System X Failure) = 1-P(System Success) = 1-0.846 = **0.154**. **System X's Risk** = (0.154)*Consequence.

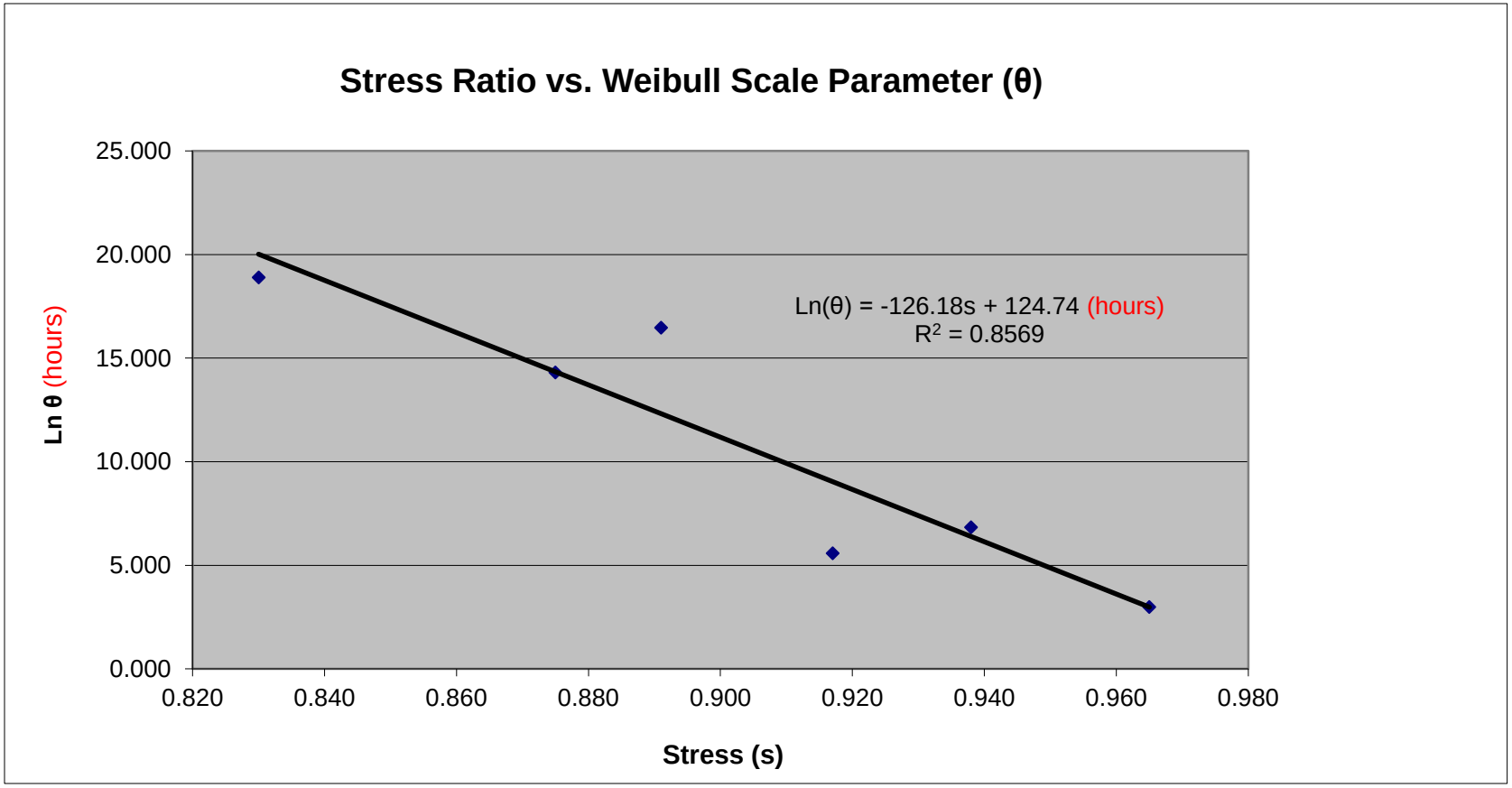
Example #6 – Based on this test data, what is the probability of success when the stress ratio is 80% and mission time is 150 days?



Example #6 – Use accelerated-life tests since test conditions vary

- It can be shown that **Ps = exp[- (t / θ) ^ β]**, the complement of the Weibull's cumulative distribution function, satisfactorily models the life data sets at each of the six stress levels (**s**). **Ps** = probability of success (as a point estimate; not interval estimate for uncertainty), **t** = mission time starting at zero, **θ** = scale parameter, and **β** = shape parameter.
- Step 1: Use the "median-rank-Y-on-X-regression" method at each stress level to determine the Weibull parameters θ and β at each stress level.
- In good accelerated-life tests, β remains relatively constant at each stress level to assure the correct failure mode. For this data, β ≈ 0.245.
- Step 2: Use the six ordered pairs (s, θ) to determine the functional relationship between the s and θ (see next page). For this data, θ* = exp[(-126.176*s)+121.563] days. θ* = 9.0387E+08 days when s = 80%.
- Answer: **Ps = 0.9784** using θ* at s = 80%, t = 150 days, and β = 0.245. Probability of failure, Pf, is 1 – Ps = 0.0216 ≈ 2 out of 100.

Step 2: Determine the relationship between s and θ



Example #7 – Obstacles for not having "zero defects"

- Classical reasons why failures occur:

Inherently Incapable: Incorrect design. Stress (load) typically exceeds strength (capacity).	**Other Time-dependent Mechanisms**: A battery runs down. Creep.
Overstressed: Applied stress exceeded the strength. Events external to design limits.	**Sneaks**: The system does not work properly even though every element (part, process) does.
Variation: Stress and strength are not fixed. Variation causes one to interfere with the other.	**Errors**: Incorrect specifications, design, software coding, assembly, test, use, or maintenance.
Wearout: Fatigue. An item is strong at the start of its life and becomes weaker with age.	**Other**:

- Knowing the potential causes of failures (defects, anomalies, loss of an item's intended function) is fundamental to preventing them.
- There are different perceptions on what kinds of events might be classified as failures. Burning O-rings on the Space Shuttle boosters were not failures until …
- An organization's reliability effort during design, development, production, operation, and service should address anticipated causes of failure as well as take in account the uncertainty involved.

Example #7 – Reliability for zero failures (discrete data)

- Time-based data and event-based data are the data types most common in determining a measure for reliability.
- **Time-based data** is the item's exposure time or run time (e.g., hours) from birth (new item) to death (failed item). **Event-based data** counts the number of failures incurred in the total number of trials or tests (n) placed on the item.
- Both classical and Bayesian statistics have methods to measure reliability when no failures occur. What follows is a method from classical statistics for event-based data.
- When event-based data (pass-fail data, Bernoulli trials) has no failures, the **table** provides the required number of consecutive successes (n) to demonstrate reliability at the level equal to the left end of a specified confidence interval. With failures, use the Clopper-Pearson interval method. Without failures, the Clopper-Pearson interval reduces to **n = ln(1-confidence)/ln(R_L)**.
- Example: An item performed 300 times with 300 successful outcomes (no failures). As per the table, the demonstrated lower-bound reliability for this process is a little better than **0.99** with a **95%** statistical confidence. Thus, the upper-bound failure probability is < 1%.

To obtain		Required trials with no failures	
Lower-Bound Reliability	Statistical Confidence	Exact Count	Rule of Thumb for Count
0.9	90%	22	2.30 x 10^#9s
0.99	90%	229	
0.999	90%	2301	
0.9999	90%	23025	
0.99999	90%	230257	
0.999999	90%	2302584	
0.9	95%	28	3.00 x 10^#9s
0.99	**95%**	**298**	
0.999	95%	2994	
0.9999	95%	29956	
0.99999	95%	299572	
0.999999	95%	2995731	
0.9	97.5%	35	3.69 x 10^#9s
0.99	97.5%	367	
0.999	97.5%	3687	
0.9999	97.5%	36887	
0.99999	97.5%	368886	
0.999999	97.5%	3688878	
0.9	99%	44	4.60 x 10^#9s
0.99	99%	458	
0.999	99%	4603	
0.9999	99%	46049	
0.99999	99%	460515	
0.999999	99%	4605168	

Example #8 – Parameter (statistical) uncertainty

- **Design Objective**: Make a survivability metric* with associated uncertainty for personnel working in the Vehicle Assembly Building (VAB) to assemble and checkout a spacecraft. This metric will combine or join two likelihoods (probabilities), namely, the likelihood of occurrence and the likelihood of impact to personnel for various hazards occurring over time, at different locations, and during different vehicle build phases.

* To provide a metric, the VAB survivability probability for each identified hazard was compared to the complement or one minus the accident rate for aerospace workers.

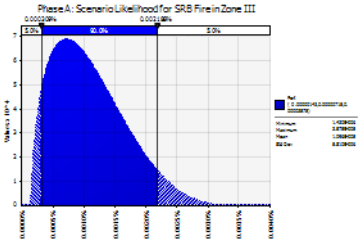
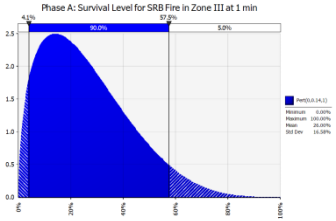
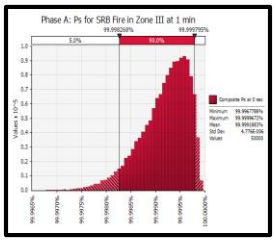
Example #8 – Analyzing risk in success space

- The survivability measure (not metric) for each scenario, at each phase, at each zone, and at each time mark was called the Probability of Survival, **P(S)**, where:
- **P(S)** = 1 - { **P(E)** * [1 - **P(S|E)**] }.
- **P(E)**, Scenario Likelihood, is the probability of the scenario occurring at any or all zones at any phase.
- **P(S|E)**, Survival Level, is the probability of survival given the hazardous event occurred in zone x and phase y. Survival Level ranges from 0% (death) to 100% (survival).
- The formulas for Aggregate Survival Level and Composite Scenarios, a group of scenarios, are not described here.

Example #8 – Input uncertainty makes output uncertainty

- If the Excel formula is: $1 - \{ P_E * [1 - P_{S|E}] \} = P_S$,

- Then the formula with Palisade's @RISK add-in to Excel makes:

$1 - \{$  $* [1 -$  $] \} =$ 

Tim.Adams@NASA.gov

Example #8 – Probabilistic simulation software

- When at least one input in the Excel equation is a probability distribution, probabilistic simulation* software such as Palisade's @RISK can be used. This software performs the following tasks:
  - Defines probability distributions for each uncertain input, the blue-shaded graphs.
  - Deterministically solves the equation numerous times (iterations) until a specified level of convergence is obtained or a fixed number of iterations were completed (e.g., 10,000 times).
  - Collects the 10,000 answers.
  - Organizes the probabilistic output (answers) into a histogram, the red-shaded graphs.
  - Converts the histogram into a probability distribution to make the area under the curve = 1.00.

* Note: Frequently, "Monte Carlo" is used to mean probabilistic simulation. **Monte Carlo** is one type of sampling method for simulation and not the only type. A common and often a preferred type of sampling method is the Latin Hypercube sampling method. Thus, it is informative when the analysts indicates the type of sampling method that was by the software to perform the probabilistic simulation.

General strategy: Thinking and producing analytically

- **COP*** is an iterative and non-linear process that logically builds ...
  - **C**oncept: talk — What concepts and data map to the desired outcome?
  - **O**peration: do — What method make the concepts and data operational?
  - **P**roduct: produce — Do the Cs and Os build, explain, and defend the P?

* Similar to Walter Shewhart's **Plan–Do–Check–Act** and W. Edwards Deming's **Plan–Do–Study–Act**.

- Expertise in the **O** task (e.g., mechanics, computation) is necessary but not sufficient to answer questions that are new to the assigned analyst.

- Thus, the manager of the analytical project can use COP as a template to inquire and distinguish between unproductive activity and results (i.e., productive tasks). Exceptions are activities that (1) "Work smarter" by researching **lessons learned** and (2) Use the **Test-Analyze-And-Fix** (TAAF) process to learn and produce iteratively.

Practice COP with the customer before the "due date"

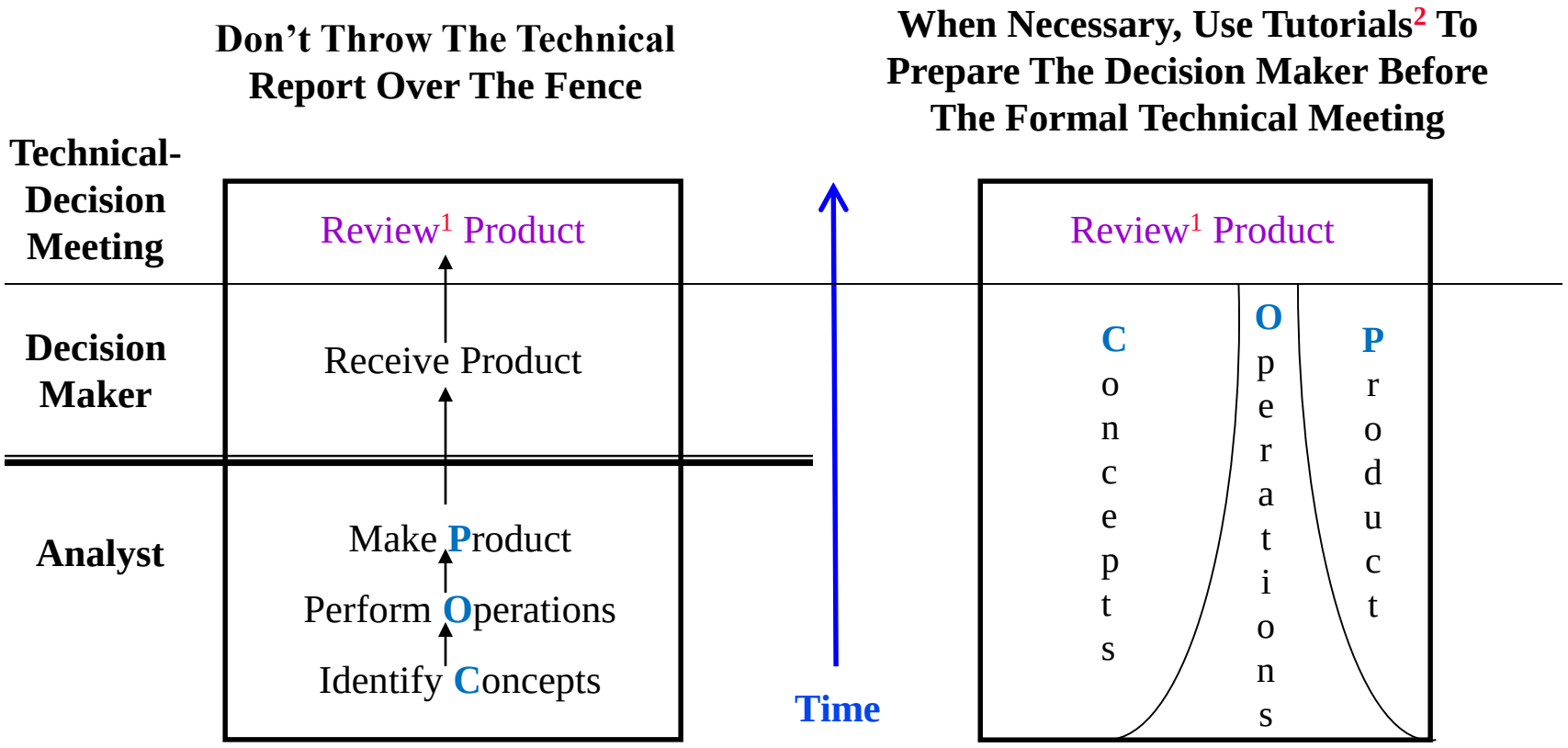
- With analytical-type work, there are advantages when the analyst is able to communicate with the decision maker while the analytical work is in the conceptual (design, thinking) and operational (build, doing) phases.

- "…only 40% of projects suggested by quantitative analysts were ever implemented. But 70% of the quantitative projects initiated by users, and fully 98% of projects suggested by top managers, were implemented."

  Barry Render & Ralph M. Stair Jr., Quantitative Analysis for Management, 6th edition

- And because … see next page.

Besides increasing the likelihood the analysis will be used, concurrent work reduces surprises and provides empathy for the decision maker



[1] Review = Decision maker(s) react positively, understand technical content as needed, and accept or reject the findings, conclusions, and recommendations.

[2] Ideally, tutorials are informal two-way discussions between the analyst and the decision maker (customer, manager) about the hope, business need, design, and development of the analysis.

Tim.Adams@NASA.gov

Analyst Mantras

- "**Success** comes in cans, not in cannots!"
  Joel H. Weldon, motivational speaker

- "**Think** about your thinking."
  The 7 Levels of Change, Rolf Smith

- "Do you think if you torture the **data** long enough it will confess to you?"
  Dr. Harold V. Huenke, Professor of Mathematics, University of Oklahoma.
  Update: Mark Hulbert's Sept 26, 2006 Market Watch stated, "If you torture the data long enough, you can get it to say just about anything."

- "Somebody is going to have to **suffer**, either the reader or the writer."
  Tom Murawski, national leader in writing improvement, The Murawski Group Inc

- "A **perfect world** is when '**big Q**' (qualitative view) and '**little q**' (quantitative view) agree--or at least understand why not to agree!"
  Author