Reliability - an Emergent System Property, 737 MAX and Societal System Examples

Jan B. Smith, PE

Key Words: Reliability, Event Interval Probability, Poisson Probability, P-values, Emergence, Reductionism, Systems Theory, Chaos Theory, Attractors, Monte Carlo, Probability Map

SUMMARY & CONCLUSIONS

A new data analysis methodology uses system level event dates to recognize unreliability of any system with complex causes and effects and forecasts the time to the next event date as a probability distribution. A single datum is often sufficient, and this makes it uniquely valuable for systems requiring catastrophic events and failures to be minimal.

The classical reliability equation $e^{-\lambda t}$ for constant failure rate λ is shown to be a systems theory emergent property revealed with event interval probability (EIP). Example datasets are technological, human, and societal to highlight the fact that the system need not be defined or even be definable. The method is shown from the classical reliability and systems theory viewpoints.

The Boeing 737 MAX crash and precursor events (extended stick-shaker activations) are two datasets showing application to technological systems. Complaints of police officer abuse of force prior to and upon the death of George Floyd in Minneapolis, Minnesota in 2020 is a human system example. A societal system example is the death of an eight-year-old child in Houston, Texas in 2020 from abuse and starvation, with the mother and another adult charged with murder. All 348 fatalities are shown to be avoidable by using EIP. A fifth dataset is what is believed to be the most referenced failure time dataset in the reliability literature. Early detection of unreliability with EIP would allow 12 of 23 equipment failures to be avoided. This has gone unrecognized despite the long history and frequent use of this dataset.

All example event interval data belong to a region of constant average event rate and the systems jump from one reliability level to the next. These reliability levels are shown to be chaos theory attractors following the universal system rule $e^{-\lambda t}$. The classical reliability equation with events being failures is a specific case of this general rule. Jumps from one reliability to another – one attractor to another – are readily seen in the datasets. The jumps to different attractors, of course, come from a change in the system of complex causes and effects, but the attractors are autonomous from the system. Causes and effects are irrelevant to EIP.

Using all data in the five examples, prior dataset values are used to forecast the next event interval. The actual realized interval to each event is compared with its forecast distribution. All datasets together provide an opportunity to make 29 forecasts as probability distributions with six of the forecast distributions being made with a single datum. All 29 realized event intervals fall within their respective 90% confidence interval, providing evidence for proof of method.

1 INTRODUCTION

Event intervals are time between events (TBE), including the time to the first event. Events can be failures, crashes, precursors, occurrences, happenings, etc., - any countable random happenstance in time. A null hypothesis is formed that event intervals are random variation from a homogeneous Poisson process (HPP) event generator with a reference, expected event rate. This null allows the Poisson distribution to be used in a reverse fashion to identify event intervals that do not fit the null hypothesis. Poisson p-values are used like conventional test statistic p-values where the null is rejected or not based upon the strength of this probability evidence. The Poisson expectation uses an event rate that may typically be a peer population mean or the mean for the dataset, along with the sum of the contiguous TBEs for that portion of the dataset used in Poisson calculations. This is further explained via example calculations. The term "event rate" and sometime "failure rate" is used when speaking in general terms; however, all calculations use the reciprocal "mean time between event(s)" (MTBE), or a variant depending upon the dataset.

Upon rejecting the null, the system generated interval(s) are used to derive a reliability probability distribution. This is also the TBE distribution. It provides the probability of the next event as a function of time. Risk is obtained by multiplying this distribution by the consequence of the event, thereby providing risk of an unreliable system remaining in service as a function of time. The math for forecasting reliability is presented and the Monte Carlo method is used to combine random variables.

A constant event rate is used in the Monte Carlo method. This assumption has been defended in conventional terms of reliability theory, engineering judgement, and experience [1]. This paper further justifies the assumption with systems theory where system reliability is shown to be a system emergent property not predicted by what is inside the system but is only determinable at the system level. Systems of all types generate random events according to a universal rule that is $e^{-\lambda t}$ or $e^{-t/MTBE}$. The classical equation for reliability with constant failure rate is only one case of this universal system rule. The constant average event rate will be shown to be a chaos theory attractor. All dataset values belong to a specific attractor. The systems and chaos theory viewpoint may seem a little mystical, but from a reliability theory viewpoint it is exactly what should be expected.

For all five datasets, there are 29 opportunities to forecast the next event interval within the attractor. Forecasts versus the actual realized intervals provide empirical evidence for proof of method. Sufficient information and data are provided to replicate the results.

2 POISSON PROBABILITY

The first step in the EIP methodology is to take a null hypothesis that the event interval data are generated by random chance from a system that produces events as an HPP – a process with events independently and identically exponentially distributed (IIED). The event rate value is a constant average and may be that of an appropriate comparative population mean or the mean for the dataset if there is history. This rate is used as a system expectation and the null is that our system is as reliable as the expectation. Inherent in this null hypothesis is that the event intervals are IIED random variables. Complex repairable equipment systems that are repaired as good as new are expected by reliability engineers to conform to the requirements of the null hypothesis. However, it is nonconforming events that are of interest. Evaluating the null hypothesis is a way to identify nonconforming events by using Poisson probability in a reverse fashion. The event intervals of interest are those that are too short or too long to be null hypothesis random chance events. Very low and very high Poisson probabilities allow the null to be rejected and the alternative hypothesis to be accepted. The alternative hypothesis is that the event interval(s) are not random chance variation of a system with an event rate equal to that of the null, but they are generated by a system with a statistically significant different event rate. Although the system elements may not have obviously changed, it has moved to a new level of reliability. The low p-values are the probability that the system is indeed reliable and that the decision to reject the null is wrong, i.e., it is the false positive.

The Poisson probability distribution of events is:

$$P(x; \mu) = (e^{-\mu})\,(\mu^{x})/x! \qquad (1)$$

Where:
$P(x; \mu)$: Probability that exactly x events occur within a specified interval when the expected number is μ
x: Number of events in an interval, x = 0, 1, 2 ….
μ: Expected number of events within the interval
e: Euler's number.

For the null hypothesis, the expected number of events, μ, over an interval is developed below and with a Poisson interval of time. A dataset must have events of the same type and have a date of occurrence. (Grouped data can be used but is not covered in this paper). The date of an occurrence is typically what distinguishes one event from another of the same type. EIP begins with these dates. The Poisson interval is often time, but time can be converted to a different Poisson interval as will be illustrated with the 737 MAX crash and precursor datasets. Below, the Poisson expectation is developed as a ratio of time between events and the average time between events.

$$\mu = \lambda t \qquad (2)$$
$$\lambda = 1/MTBE$$
$$t = \sum TBE$$
$$\text{therefore:} \quad \mu = \sum TBE/MTBE \qquad (3)$$

Where:
λ: Expected event rate, event counts per time unit
t: Time between event (TBE) or $\sum$TBE
MTBE: Expected (Mean) time between events
$\sum$TBE: sum of time between appropriate contiguous events

In this paper event rate λ is used in general description, as it seems more familiar than using MTBE. For calculations, however, the complement of event rate, MTBE, is always used.

3 BOEING 737 MAX CRASH INTERVALS

For evaluation of aircraft crash events, departures are the better Poisson interval with departures between event (DBE) and its mean (MDBE) used for TBE and MTBE, respectively. Also, the peer group mean is in units of departures. For the null hypothesis to which we apply the above equations, there is no distinction between the first event interval and subsequent intervals. In some data analysis contexts, departures to first event would be appropriate, but here we can use the term "departures between event" for even the first event.

3.1 *First crash*

The first crash on 10/29/2018 occurred with a 737 MAX fleet DBE of 135,980 departures [2]. The worldwide jet commercial scheduled carrier fleet is a peer population against which the 737 MAX fleet is compared. This peer population MDBE is 6,105,714 calculated from published data [3] - number of fatal crashes over a recent 10-year period and the number of departures over the same period. With Poisson probability for the first event being defined as event 1 p-v1, the probability of one or more events within the interval of the one event is obtained using the complement of equation 1.

$$\text{p-v1} = 1- (e^{-\mu})\,(\mu^{x})/x! = 0.02202$$

where:
x=0
DBE = 135,980
MDBE = 6,105,714
μ = DBE/MDBE = 135,980/6,105,714 = 0.02271

Event 1 p-v1 answers the question, what is the probability of one or more events occurring within the interval of the one event? This Poisson probability is used like a test statistic p-value as strength of evidence against the null. This low p-value indicates how poorly our data fits the null hypothesis. The data is not likely to be random variation in a fleet that is as reliable as its peer group, the worldwide fleet. This p-value is sufficiently low to reject the null and accept the alternative hypothesis that the 737 MAX fleet is unreliable compared to the peer group.

3.2 *Second crash*

The number of departures to the second crash is not believed to be publicly available and is estimated from data that are available. Plane delivery dates are available from the Boeing website. The delivery date for individual planes allows plane-days in service to be aggregated for the fleet. The first crash plane-days in service is 47,711 and to the next is 41,352. From the departures to the first crash and the plane-days to the first crash, the average number of departures per day is 2.85. Using this first interval average as an estimate for the second interval gives 2.85 * 41,352 = 117,856 departures.

Using equation 1 and 3, the second event p-v1 (the probability of the second event occurring in an interval of the second event) is:

$$p\text{-}v1 = 1 - (e^{-\mu}) (\mu^{x})/x!) = 0.01912$$

where:
x=0
DBE = 117,856
MDBE = 6,105,714
μ = 117,856/6,105,714 = 0.01930

Using equation 1 and 3 with departures as the Poisson interval, the second event p-v2, the probability of two or more events within the interval of the two events, is:

$$p\text{-}v2 = 1 - \{(e^{-\mu}) (\mu^{x0})/x_0!) + (e^{-\mu}) (\mu^{x1})/x_1!)\} = 0.00084$$

where:
$x_0 = 0$
$x_1 = 1$
$\mu = \sum$DBE/MDBE = (135,980 + 117856)/ 6,105,714 = 253,836/6,105,714 = 0.04280
MDBE = 6,105,714

The probability of two or more crashes within the interval of the two crashes, if the fleet is as reliable as the peer group, is 0.00084. The approximately eight chances in 10,000 that the 737 MAX is as reliable as the worldwide fleet is overwhelming evidence for rejecting the null. The alternative hypothesis that the fleet is less reliable than its peers could be accepted after the first crash, and certainly after the second. But it was not, and unreliable aircraft continued in service for three days after the second crash before the fleet was grounded at a risk of approximately 7.9 additional fatalities on average [1].

Figure 1 is the upper portion of a Microsoft Excel spreadsheet designed to easily implement EIP analyses. The equations used for p-v1 and p-v2 are in the appropriate cells for calculation using DBE, in green, and MDBE input data. (Sufficient information for constructing the workbook is in earlier papers [4] and is available without cost at www.pmfseries.com). What is the meaning of the probability values? They are the probabilities that the system is as reliable as the reference system, in this case the worldwide fleet. If the null is rejected, they are the probabilities of the decision being wrong. They are the probabilities that the worldwide fleet could have intervals that short by random chance. They are the false positive values.

Event	date	Event #	DBE	MDBE	p-v1	p-v2
Lion Air	10/29/2018	1	135,980	6,105,714	0.02202	
Ethiopian Airlines	3/10/2019	2	117,856	6,105,714	0.01912	0.00084

Figure 1- 737 MAX departures between crashes in green with worldwide mean placed in a Poisson probability map. Equations for p-v1 and p-v2 (as just reviewed) calculate appropriate probability values in the right most columns of the spreadsheet. Color formatting brings visual attention to actionable p-values.

4 *737 MAX PRECURSOR INTERVALS*

EIP application to precursor events are applied to the 737 MAX to indicate how the first crash could have been avoided; thereby, illustrating how to avoid similar events in the future. Precursors are less severe events that can signal a system problem. The first precursor “extended stick-shaker activation” occurred on the first fatal crash plane on the prior flight on 10/28/2019. The cause of the precursor was not corrected, nor was the affected plane removed from service for investigation. The issue continued the next flight – the fatal flight.

A second extended stick-shaker activation occurred on the second fatal flight. Probability values for the two instances are seen in figure 2. (Reference 1 published in 2023 updates prior papers on the Boeing 737 MAX event interval probability input data and results. It also provides additional information on data sources, analysis considerations and details, and a more thorough discussion of EIP specifically applied to the 737 MAX. This paper covers only the basics for the 737 MAX analysis).

We see that both crash intervals and precursor event intervals are sufficiently low to reject the null hypothesis and accept that the 737 MAX fleet is unreliable relative to the peer population upon the first event, and overwhelmingly so upon the second event. The next step in EIP is to determine the actual reliability from the system performance data and forecast future failures as a probability distribution. But before this next forecasting step, the analysis method is applied to social and human systems to demonstrate how the method is applicable to any event from a complex system of any type.

5 *CPS ABUSE INVESTIGATIONS*

In year 2020, a Houston, Texas, 8-year-old child died by abuse and starvation. The mother and her domestic partner were charged with murder. This case study is important for engineers to consider because it demonstrates EIP applicability

event	date	Event #	Plane-days between event	MP-dBE	p-v1	p-v2
Lion Air	10/28/2018	1	47,481	1,996,497	0.02350	
Ethiopian Airlines	3/10/2019	2	41,582	1,996,497	0.02061	0.00097

Figure 2 – Extended stick-shaker activations are crash precursors. The number of these events for all of Boeing’s fleets in a specific period and fleet plane-days in-service over the same period, obtained from plane delivery dates on Boeing’s website, provide the Poisson expectation μ against which the 737 MAX precursors can be compared [1].

to events in general and complex systems in general. Experts in particular systems and events, that we will call subject matter experts (SMEs), are commonly resistant to accepting EIP as a reality apart from what is inside the system. Use of EIP in non-technological applications may improve their understanding and acceptance of system performance as an independent measurement of reliability that is independent of what is inside the system. The system that generates Texas Family and Protective Services (that we will call the more general Child Protective Services or CPS) abuse investigations is both undefined and undefinable. Without any knowledge of the causes and effects within the system, the reliability of the system can be completely understood. Later, we explain that reliability of complex systems is a systems theory emergent property.

Incidences of abuse are hidden. They are unknown and unknowable. The system that produces an abusive environment is a subsystem within our system. CPS investigations are triggered by reports of abuse or possible abuse. This reporting system is also a subsystem within our system. These two subsystems are both complex and work together to produce a CPS investigation. CPS investigation dates are all we have available and are all that is needed for EIP, as with the 737 MAX datasets. Due to the death, CPS investigation dates for this case are publicly available [5]. The mean days between child abuse in Texas is available from the number of children in Texas and the number of investigations per year [6]. The mean is used for the Poisson expectation in equations 1 and 3. Figure 3 is the probability map that organizes the equations, allows efficient data analysis and trends probability values. Sufficient Excel equations are shown to allow all the equations to be understood upon inspection of the patterns. Excel rows and columns are shown. The equations for p-v1 and p-v2 were shown for the 737 MAX crash dataset. To explain an additional p-value for understanding, the equation in cell M7 is for p-v3 at the 4th event on 12/16/2020. The cell value answers the question "what is the probability of three or more investigations within the sum of the intervals of the last three investigations"? Embedded Excel formulas are used allowing the most efficient answer to our question to be the complement of the cumulative probability of two events. The probability value of 0.00007 means there are seven chances out of 100,000 that there could be three or more events within the time of these three events for an average Texas child. This rare chance rejects the null hypothesis that the investigation dates are random variation of a normal child's "system" of abuse and abuse reporting. Armed with this information, investigative attention could be focused on this particular child. If EIP had been used to recognize the significance of investigation dates, the death of the child five days later could have been prevented. The last event is not the normal CPS investigation, but it is the death of the child. The 5-day interval is shorter than EIP prediction. It is statistically different from prior intervals and is statistical evidence that the child was dying from abuse and starvation upon the 12/16/2020 investigation when, because of the Covid-19 pandemic, the child was not observed. The mother and her domestic partner were charged with murder by abuse and starvation. Early identification of the child's probabilistic data could have avoided not only death but 594 days of avoidable exposure to abuse and starvation. In the United States alone, five children per day die by abuse [7]. Applying science to event intervals could save some by informing stakeholders early.

6 *POLICE USE OF FORCE*

Derek Chauvin, the former Minneapolis, Minnesota police office found guilty in the death of George Floyd by use-of-force in 2020, was identifiable as unreliable 994 days prior with 0.00199 probability of a false positive (figure 4 event 4 p-v3). At 465 days prior to Floyd's death, the probability of the system being as reliability as the reference was 0.00064 (figure 4 event 5 p-v4); therefore, the system was most assuredly unreliable. It should be obvious without explanation how early identification of police officers that are statistically prone to use of force would have a favorable impact on American society by forcing a more stringent police complaint investigation on those that are statistical outliers. The Chauvin system illustrates that calculations involve no knowledge of the system that generates occasions of use-of-force. Only the dates on which our system generates an event are used. The event intervals determined by citizen complaint dates [8] and the mean complaint rate for police officers in large municipalities [9] are used in the calculations. The system is everything that influenced the dates. The system does not need to be defined or understood. The

A	B	C	D	E	F	K	L	M	N	O	P
2	Event	date	Event #	DBE	MDBE	p-v1	p-v2	p-v3	p-v4	p-v5	p-v6
3	approx birthdate	6/21/2012									
4	CPS investigation	10/1/2018	1	2,293	10,564	0.19512					
5	CPS investigation	5/7/2019	2	218	10,564	0.02043	0.02415				
6	CPS investigation	8/26/2019	3	111	10,564	0.01045	0.00048	0.00212			
7	CPS investigation	12/16/2020	4	478	10,564	0.04424	0.00150	0.00007	0.00024		
8	death/murder charges	12/21/2020	5	5	10,564	0.00047	0.00101	0.00003	0.00000	0.00001	

=(1-EXP(-E5/F5)*POWER(E5/F5,0))/FACT(0)

=1-POISSON.DIST(2,SUM(E5:E7)/F7,TRUE)

Figure 3 – CPS investigations of an 8-year-old child. DBE is days between events and MDBE is mean days between events. The first event on 10/1/2018 was not statistically significant. Events 2, 3 and 4 show increasing strength of evidence that intervals are not IIED with mean of 10,564 days (population / investigations), the peer group average. Excel rows and columns are shown so that equations for probability on the right can be read. The few equations shown reveal the pattern for all probability equations. The last event is the child's death by abuse and starvation that was preventable by recognizing the increasing probability evidence for abuse that could allow investigative resources to be more focused according to the individual child's risk.

Event	Date	Event #	TBE	MTBE	p-v1	p-v2	p-v3	p-v4	p-v5	p-v6
service year	**1/1/2001**									
Julian Hernandaz	**2/15/2015**	1	5158	3842	0.73881					
Jimmy Bostic	**4/15/2016**	2	425	3842	0.10472	0.42636				
Zoya Code	**6/25/2017**	3	436	3842	0.10728	0.02166	0.20804			
John Pope	**9/4/2017**	4	71	3842	0.01831	0.00798	0.00199	0.07677		
Sir Riley Peet	**2/15/2019**	5	529	3842	0.12863	0.01100	0.00267	0.00064	0.03108	
George Floyd	**5/25/2020**	6	465	3842	0.11399	0.02822	0.00289	0.00071	0.00017	0.01156

Figure 4 – Police use of force complaints against officer Chauvin upon and prior to the death of Floyd. Poisson probability values overwhelming reject the null hypothesis 994 days before the last event. These probabilities and their trends could inform any stakeholder and lead to more focused and thorough complaint investigations.

probability map is figure 4. (There is evidence that data are missing prior to 2/15/2015, but that is not essential to the purpose of this paper. Also, it would only lower the p-values).

7 RELIABILITY DISTRIBUTIONS

Time to next event probability distributions are obtained with Monte Carlo simulation. Monte Carlo is used to solve the math problem of adding and multiplying random variables – adding and multiplying probability distributions. Using Monte Carlo may seem too complex to be practical, but the equations to be reviewed are easy to manage with computer spreadsheets with equations positioned to achieve the purpose. (The workbook can be built by the user or downloaded at no cost). The distributions will provide the probability of the next event occurring as a function of time. Time is converted to departures between events and plane-days between events for crashes and precursors, respectively. Risk of continued system operation, without any effectual change, is found by the product of this probability and the consequences of an event. This is especially critical if the system has been discovered to be unreliable relative to expectation.

Figure 5 graphically demonstrates the basic process for Monte Carlo simulation to obtain samples of time to the next failure (TBF) where the events are failures. Many random number draws, or iterations, form probability distributions for TBF. The cumulative failure distribution, equation 4, is the complement of reliability, or unreliability. This is the complement of equation 1 with x = 0. Equation 4 is well known and accepted by reliability engineers, so this will be used; however, we could just as well start with the equation for reliability and get a mirror image of figure 5 that contains identical information. The math applies to all systems that behave as complex repairable systems. The 737 MAX is such a system, although it is not repairable in the way one may think. The crashed planes are simply removed from service and the total planes in the fleet are reduced by one then two, and the repair time is zero. The math for the distributions applies to any system of any kind when the causes and effects generating the event are complex, i.e., events are IIED random variables. This allows use of a constant failure rate.

$$F(t) = 1-e^{-\lambda t} \quad (4)$$

Failure rate λ in equations 1 and 4 is a constant average. Before the null is rejected, the null hypothesis assures the constant failure rate by definition – it is inherent to the null hypothesis. Now that the null is rejected, use of equations 1 and 4 requires an assumption of a constant failure rate. This is expected to be the case for complex repairable systems and is always the case for systems that behave as complex repairable systems. The constant average event rate is justified from reliability theory and experience [1]. It will be further justified from a systems theory perspective later in this paper.

A uniformly distributed zero to one random number is transformed to a time sample using Equation 4. Equation 4 is set equal to a random number (RN).

$F(t) = 1-e^{-\lambda t} = RN$
Where:
RN = uniformly distributed random number from 0 to 1
λ = failure rate or, more generally, event rate
t = time

rearranging terms,

$e^{-\lambda t} = 1- RN$

Event rate is constant, so λ = 1/MTBE and the complement of a 0 to 1 uniformly distributed random number is a 0 to 1 uniformly distributed random number, so

$$e^{-t/MTBE} = RN \quad (5)$$

Taking natural log of both sides and simplifying,

$\ln e^{-t/MTBE} = \ln(RN)$
$-t/MTBE = \ln(RN)$
$t = -MTBE*(\ln(RN)) = TBE\ sample$

The time to first event sometimes must be distinguished from subsequent time between events, but for constant λ the term TBE can also be used for the first event. So, equation 6 is for the length of any event interval, including the first:

$$TBE = -MTBE*(\ln(RN)) \quad (6)$$

MTBE in equation 6 can be a population mean that is single-valued, as discussed thus far, and produce the blue

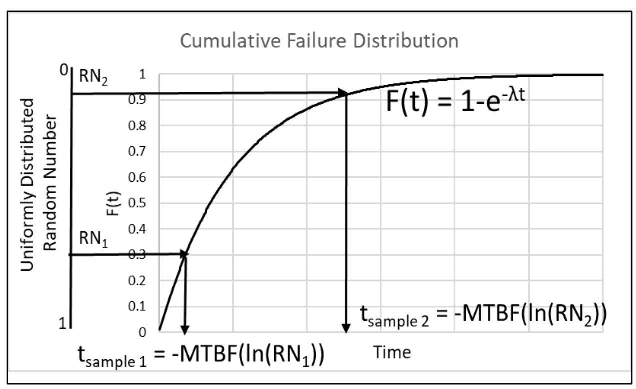

Figure 5 – An event interval sample, here called failure time sample, is determined with a random number draw from the secondary y-axis that is transformed by equation 4 to a failure time. The graph shows two samples of failure times derived by equation 6.

distribution in figure 6. Without a peer group, the running average is used and is a sample statistic from a population of unknow mean. After the null is rejected, the data are always a sample from a system of unknown mean. MTBE as a random variable is explained in the context of forecasting the next event.

The number of TBE values available for the next event forecast is n, where n can be any positive integer including only one. Notice that the number of events available is used in the forecast. A restraining paradigm is that using only a little data in the forecast means the forecast cannot be accurate. In our examples, getting more data means having more fatalities. Data analysis methods are needed that use ultra small datasets, especially when the data are catastrophic events. The method following applies to any number of events starting at one.

Starting with a TBE dataset used to forecast the next event interval as a probability distribution, equation 6 is used to generate TBE samples using the dataset mean – sample mean – to get a sample of the population mean as a random variable.

$$MTBE_{RVs1} = (TBE_{s1} + TBE_{s2} + \ldots\ldots TBE_{sn})/n$$

Where: $MTBE_{RVs1}$ = 1st sample of the mean TBE as a random variable
TBE_{sn} = Time between event random variable sample n
n = number of TBE values used for the forecast

Using equation 6 with the dataset mean,

$$MTBE_{RVs1} = \{[-MTBE_D*\ln(RN_{s1})] + [-MTBE_D*\ln(RN_{s2})] + \ldots\ldots[-MTBE_D*\ln(RN_{sn})\}/n \quad (7)$$

Where: $MTBE_D$ = dataset mean
RN_{sn} = Random number draw, one for each n

The RN subscripts above are to assure that each random number is independent of any other draw. When the random numbers are assured to be independent during the Monte Carlo simulation – such as when a computer spreadsheet is properly designed - we can omit the RN subscript. Now the MTBE random variable sample is used in equation 6 a second time with the mean as a random variable from equation 7. But this is all within the same calculation – the same Monte Carlo iteration – for equation 8.

$$TBE_{Fs} = -MTBE_{RVs1}*\ln(RN_{sn+1}) \quad (8)$$

Where: TBE_{Fs} = forecast sample, sample of TBE
RN_{sn+1} = an additional unique RN.

Figure 6 is 100,000 iterations or samples of TBE_{Fs} that are sorted from smallest to largest to form a TBE distribution for the 737 MAX second crash using the number of departures to the first crash as input data. This distribution is the forecast for a second crash. The actual experienced departures between the first and second crash is 117,856, by calculated estimate. This number of intervals projected to the cumulative probability y-axis is expected to not be exceeded with probability 0.69. In other words, two thirds of the time the second crash would occur even sooner than it did.

Figure 6 for departures to the next crash should be viewed as math solutions and not as computer simulations in the normal sense. Monte Carlo generated random numbers is an efficient way to perform this convolution of two probability distributions. The resulting distributions accurately reflect uncertainty of outcome, but because the addition and multiplication of probability distributions are by Monte Carlo, the accurate results have imprecision. To illustrate the imprecision of the accurate forecast, the process for generating the second crash forecast distribution of figure 6 was executed five times. The actual interval of 117,856 projected to cumulative probability with five simulations of 100,000 iterations produced probability results ranging from 0.68737 to 0.68927.

An airworthiness directive (AD) was issued by the FAA nine days after the first crash. Even if the AD totally resolved the problem, the 9-day delay incurred significant risk that is currently unrecognized. There were 227 planes in service during the nine days with an average of 2.85 departures per day, as calculated earlier, resulting in about 5,823 departures in the period. Using data that formed figure 6 second crash forecast distribution, the probability of a crash within nine days is about 0.133. The average number of fatalities can be considered the consequence of a crash, or 346/2 = 173 fatalities/crash. Risk is probability times consequence or 0.133 * 173 = 23.0 fatalities on average.

Similarly, following the second crash there was a 3-day delay in grounding the fleet. The fatality risk for these 3-days, considering all the fleet to have been flying, is about 7.9 fatalities [1]. These risks obviously would not have been taken if there was awareness.

The reader may at this point think that Boeing and the FAA acted inappropriately. But consider that everyone, including academia and this author, has overlooked the information found in only one or two event dates as there was no analysis method to reveal the information.

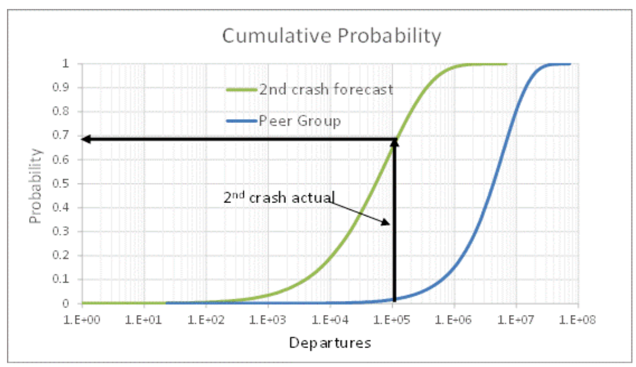

Figure 6 – Forecast for a second 737 MAX crash, in green, using the first crash interval as in equations 7 and 8. This figure is like figure 5, but with log scaling on the x-axis. The blue distribution is the worldwide fleet DBE. The realized 2nd crash interval projected to the cumulative probability axis of its forecast shows 0.69 probability of occurring prior to the actual interval.

8 SYSTEMS THEORY

As explained by Hitchens [10], some complex systems can be better described by their behavior than by delving into the depths of system design, structures, processes, etc., and this provides a way to manage complexity. Systems exhibit emergent properties that cannot necessarily be explained or predicted by understanding the complexity of components and the interactions among themselves and their environment. Hitchens describes rare universal rules that apply to systems. Such a universal rule controls each individual datum in our five datasets – 37 in total. The rule is equation 5, for systems with a constant event rate. Systems with a complexity of causes and effects are expected to have a constant event rate from classical reliability. The $e^{-\lambda t}$ or $e^{-t/MTBE}$ rule is followed for all event intervals generated from any system that behaves as a complex system, defined as one in which events are generated by a complexity of causes and effects. This is the equation for reliability of a constant failure rate system. A systems theory perspective allows reliability to be seen as an emergent property of the system. The math is identical to the classical reliability perspective, as that perspective is a specific case of the more general system perspective. The system does not need to be defined as the event intervals are autonomous from what is inside the system; furthermore, EIP analysis using this autonomous data is independent from what is inside the system – it has nothing to do with what is inside the system. This fact is, thus far and in general, incomprehensible to system subject matter experts (SMEs) and reliability professionals as is discussed in the next section.

The independence of EIP and system contents is demonstrated by the child abuse example. The system that generates CPS investigation events contains two subsystems. One subsystem produces an abusive environment for the child and the other is an abuse reporting subsystem. Both subsystems are complex and unknown, and so much more are the interactions between these subsystems to produce the reports that trigger CPS investigations.

The four datasets discussed thus far either begin with or move to a constant event rate. Also, a fifth dataset seen in figure 7, shows two distinct constant failure rates and the system jumps from one to another. Figure 7 trends the 60-year-old dataset for an airplane air conditioner. This dataset is thought to be the most referenced failure time dataset in the reliability literature [11], yet there is an unrecognized phenomenon. TBE data residuals are trended. This method for trending event data was developed for maximum sensitivity to changes in event rate, reliability, and availability. The dataset shows two distinctly different system reliabilities for the same equipment. Although the system is physically the same throughout the dataset, system reliability jumps back and forth between two reliabilities. Emergent reliability is attracted to one of two regions reflecting a change in the parameter MTBE in equation 5, where chaos, or unpredictably, might otherwise be expected. In figure 7, these two regions are called reliable and unreliable with the unreliable regions enclosed within ellipses. In the language of chaos theory, these regions are attractors.

How randomness in complex systems translates into events in time is seen in figure 5. The 0 to1 uniformly distributed random number on the secondary y-axis represents chaos that the universal system rule in equation 5 translates into events in time on the x-axis. Dataset intervals are used to first check if they are consistent with the null hypothesis curve shape. When the null is rejected, the dataset is used to determine the actual shape of the curve using Monte Carlo.

Viewing reliability as an emergent property as revealed by EIP allows us to know all about system reliability even when knowing little or nothing about the system generating the events. We can imagine the data of figure 7 coming from a signal received from outer space. This signal data is generated from some deterministic cause and effect process that we call a system. The system could be technological, natural, social, or any combination. We may have no idea how the signal is generated or even what the signal means, but we know the two emergent reliability levels of the signal generating system and the probability distribution for the next signal within that attractor. (In the general case, system reliability is an example of strong emergence, though many scientists deny the existence of such. Weak emergence, accepted by scientists, requires the system to be definable so that, at least in principle, the emergent property can be deduced from system contents; however, in the general case, system contents are scientifically unknowable. Discussion is beyond the scope of this paper).

All the datasets have regions of constant average event rate where the $e^{-t/MTBE}$ universal rule shows a certain reliability level. This is additional support for the constant event rate used to develop probability distributions for forecasting the next event. For all five datasets, there are 29 opportunities to forecast the next event using the prior events in that attractor. Six of the 29 forecasts used a single data point - a single date. The actual, experienced event interval is compared with the interval's forecast distribution. This is done in the manner shown in figure 6 where the second crash actual interval is projected from the x-axis via the probability distribution to the cumulative probability y-axis where 0.69 probability is read. This was done for all 29 forecasts with results in figure 8. The actual realized intervals all fell well within their respective forecast and, as well, they were scattered randomly as they should be as seen in figure 8. This is empirical evidence for proof of method. All data necessary to replicate the results are in this paper.

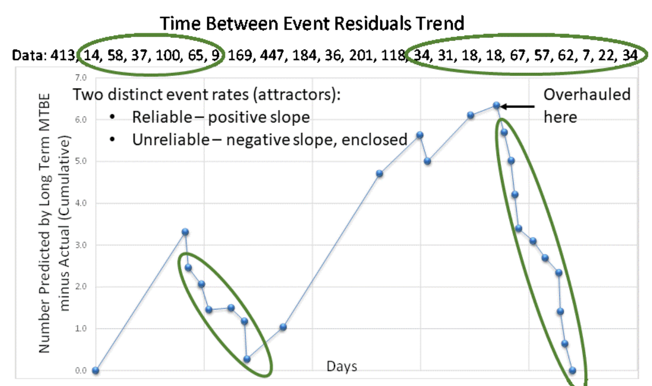

Figure 7 – A residuals trend for viewing change sensitively. Graph construction is revealed by axis description. There are two stable regions, and the system jumps back and forth between the two. The one labeled "unreliable" is enclosed with an ellipse. An overhaul at the 13th event introduced a change to which the system is sensitive. Something similar happened upon events 1 and 7 repairs causing the system to jump attractors. Such information is needed in root cause analysis. 70% of the failures are in the unreliable attractor.

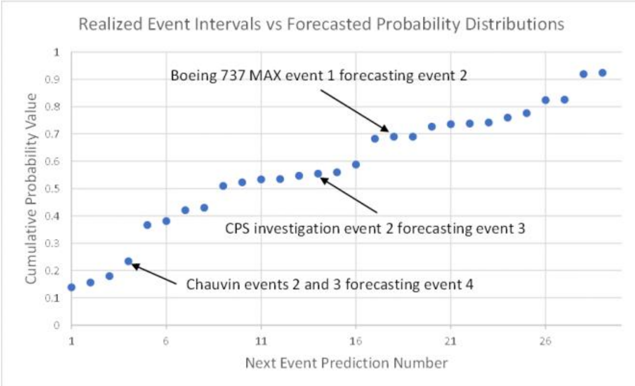

Figure 8 – 29 realized events versus their respective forecast as in figure 6 plotted smallest to largest. An event and any prior events in the same attractor were used to forecast the next event interval cumulative probability distribution. All actual results are within their respective 90 % confidence interval. Composite results are randomly scattered with 0.576 average versus 0.632 ideal. Realized results that match forecasted is empirical proof of method.

9 REDUCTIONISM VERSUS EMERGENCE

Reductionism is on the opposite side of the spectrum from emergence. Reductionism is going inside the system to reduce the complexity of parts and their interactions to smaller elements. If specific causes of unreliability are to be understood and corrected, it will be through reductionism. Engineers and scientists are taught by both education and experience to solve problems by breaking them down into smaller, manageable pieces. Use of system performance for reliability (emergence) seems incomplete and insufficient to any SME because they tend to be reductionist, whether reliability or aerospace engineer or, this engineer supposes, even sociologist. There is tension between reductionism and emergence because, according to Damper [12] "the two creeds seem to work in opposite directions." These two divergent views (systems of thought, philosophies) regarding systems are often seen to be mutually exclusive with arguments as to which is true. This is evident in Damper's use of the word "creed" in his description of these two views. Reductionists and emergentists seem to believe that their own philosophy precludes the other view. But both views are valid in context and need to be taken one at a time. Emergence and EIP should be considered first to see if the event could be random variation of a reliable system or evidence of a jump to a different and unacceptable attractor.

10 RECOMMENDATIONS

The FAA should use EIP as a second layer of safety that is independent of the design and certification process. Events to be analyzed should obviously include major events such as crashes and plane-related emergency landings as in the first dataset, but precursors should also be sought out and the significance of these event intervals considered as in the second dataset.

NTSB crash investigations should include any missed opportunities to avoid a crash, such as those described in the first two datasets.

REFERENCES

1. J. Smith, "Avoiding System Failures with Event Interval Probability – 737 MAX Case Study", *Proc. Ann. Reliability & Maintainability Symp*., Jan 2023.
2. Lewis Lamb contributor (Seattle Times), DocumentCloud, 6573647-TARAM-for-MAX.
3. Boeing, "Statistical Summary of Commercial Jet Airplane Accidents Worldwide Operations," 2018.
4. J. Smith, "Failure Interval Probabilistic Analysis for Risk-based Decisions – Concorde Crash Example, *Journal of System Safety*, Vol. 56 No. 3 Spring 2021.
5. Texas Department of Family Protection Services, *Child Fatality Report Release of Information to the Public*, Mar. 23, 2022.
6. Texas Department of Family and Protective Services, Child Protective Investigations (CPI), online Data Book, 2021 all Texas data.
7. Child Welfare Information Gateway. (2021). *Child abuse and neglect fatalities 2019: Statistics and interventions*. U.S. Department of Health and Human Services, Administration for Children and Families, Children's Bureau, p 2.
8. A. Vansickle, J. Lartey, published in partnership with New York Times, "That Could Have Been Me," Feb. 2, 2021.
9. U.S. Department of Justice, Department of Justice Statistics, Citizen Complaints about Police Use of Force, June 2006.
10. D. K. Hitchins, Systems Engineering, West Sussex, England, John Wiley & Sons, 2007., pp 11-27.
11. H. Ascher, H. Feingold, Repairable Systems Reliability – modeling, inference, misconceptions and their causes, New York, Marcel Dekker, 1984, pp 144-145.
12. R. I. Damper, Editorial for the Special Issue on 'Emergent Properties of Complex Systems': Emergence and levels of abstraction, International Journal of Systems Science, Vol. 31, no. 7, 2000, p 816,

BIOGRAPHY

Jan B. Smith, PE
3342 Courtland Manor Ln.
Kingwood, Texas, USA
e-mail: jansmith@pmfseries.com

Jan Smith, PE, BSME, reliability engineering consultant with a 58-year career in reliability engineering, established and managed reliability departments. He established two professional engineering firms specializing in reliability engineering, including within a major corporation (now Honeywell), for which he was the responsible engineer. Now an independent consultant, he is focused on integrating new system level concepts for reliability and availability into existing decision-making processes, with automation of the data analysis when volume or speed dictate. A registered PE, he holds three patents using system level data for probabilistic availability and developed and applied automated EIP analysis for asset reliability. Website: https://www.pmfseries.com.