Failure Time Analysis Applied to Boeing 737 MAX

Jan B. Smith, PE, Zachry Industrial Inc.

Key Words: Homogeneous Poisson Process, Poisson Probability, Repairable Systems, Monte Carlo, Reliability, Simulation

SUMMARY & CONCLUSIONS

Data for the analysis comes from Boeing publications and public announcements. The analysis input data are believed to be conservative, i.e., the data produces results that tend to understate the significance of the analysis findings and conclusions. Using data existing at the time, the analysis shows the 737 MAX should have been grounded sooner. Perhaps after the first event (crash) but certainly after the second. The decision not to ground after the second event led to high risk of a third event. If all planes remained in service, the risk of a third event in three days is about 1 in 25. This high risk indicates that Boeing and the FAA did not do a failure time analysis. It is supposed those that chose to voluntary ground their aircraft made the decision without data analysis as well. At the time of the events, cause and effect data were sparse, but the event time data were readily and immediately available. This unused data were of tremendous value in quantifying risk for decision making. An objective of this paper is to assure failure time data are analyzed in the future, not only in engineered systems, but in business, medical, and any other data system where special cause events are of interest and a homogeneous Poisson process (HPP) is a reasonable null hypothesis.

The analysis is shown in sufficient detail so Boeing, the FAA or accident investigators can conduct an analysis with precise input data. The analysis method is of general interest broader than the two crash investigations. They are of course obligated to correct or improve the analysis methodology as well. Human nature will tend to blame Boeing and the FAA for not conducting the analysis for their risk related decisions regarding grounding. While understandable, this would be irrational. Boeing and the FAA, like essentially everyone, did not know to do the analysis. The Poisson methodology was only published in early 2018 [1]. It is otherwise not in any text books and technical papers that the author has encountered. A different concept, Monte Carlo simulation, confirms Poisson results. (These concepts are completely different but share a HPP null hypothesis - explicit with Poisson and implicit with simulation). Both concepts are simple and seemingly obvious once recognized. Accident investigators and others should consider if it is possible to hold Boeing and the FAA accountable for not doing a particular data analysis that data analysis experts don't do.

The 737 MAX example shows how Poisson and simulation failure time analysis can use even a single event for powerful information useful in decision-making. It also illustrates how we are essentially always too late in doing the analysis. For in-service repairable systems, (and in general any point process for which a HPP null hypothesis is reasonable) the primary value of analyzing intervals is to identify events that may be special cause events needing immediate intervention. The value is in immediate intervention and avoidance of otherwise predestined failures. But a primary restraining paradigm is that failure time analysis is used only when a need is otherwise identified. So we don't use failure time analysis when we should, i.e., before we know that we need to use it. We only use it when it is too late to be of greatest value, i.e., after we know we need to use it. The result is we derive lower value in the analysis and this reinforces the paradigm. If the analysis is of low value, then there is little justification in doing it, especially when there is no obvious need. It may be beyond the scope of accident investigators (and certainly beyond the scope of this paper), but it is of general interest to study the restraining paradigms that culminated in the absence of failure time analysis on the 737 MAX. The mathematics for Poisson have been around for nearly 200 years. Computer simulation has been readily available for over 30 years. Why aren't we using them?

How can an analysis be done prior to recognizing a need? Computers can automatically analyze all failure times immediately upon failure and report only those that are statistically significant. Perhaps ultra-critical events like fatal accidents can be analyzed contemporaneously upon occasion by policy and procedure. But most applications will require automated data analysis to find the few events of interest among numerous events on numerous assets. With much data being electronic, automation is practical, but this requires a major shift to a new paradigm in how, when, and why we use failure time analysis.

Poisson methodology automation using Microsoft Excel has been demonstrated and is being applied in a process industry plant. About 2,500 work orders per year are automatically and contemporaneously analyzed with a series of Poisson p-values that flag special cause events for intervention. Once the automation identifies low p-value events, data may need editing (the data are dirty) and analyzed as a single dataset. Single datasets are easily analyzed one at a time within Microsoft Excel with tables called probability maps, with confidence intervals determined with commercially available Excel add-in packages for Monte Carlo simulation. Cell equations are given so probability maps of any size can be built. Automated analysis needs commercial software providers to transition from the current "one dataset at a time" analysis capability to the functionality of proactively screening large

numbers of individual datasets to identify the few of interest.

1 INTRODUCTION

This paper is specifically directed to failure time analysis of in-service repairable equipment and systems such as found in process plants. In process plants, it is important to identify special cause failures immediately to provide the opportunity to intervene in the repair planning and execution. By recognizing reliability degradation at an opportune time, intervention can avoid future failures. It is uncommon to use Poisson probability for failure time analysis, but it is required to detect failure rate step change. Coincidentally, the Boeing 737 MAX is an in-service repairable system and provides an example of useful decision information obtainable with only a few failures. The 737 MAX also exemplifies how failure time analysis is invariably done too late to be of maximum value.

2 POISSON PROBABILITY

The Poisson distribution is used to determine the probability of specific numbers of events occurring within a specified time interval, when the events are generated by a HPP. Failures times are independent and identically distributed exponential random variables. The mean number of events must be constant for any time interval of equal length. Repairable system failures are, in general, such a HPP. But new failure modes, improper repair, and any other special cause produces time-between-failure (TBF) data that do not fit the HPP conditions for Poisson. Moreover, it is these nonconforming special cause failures that are of most interest. Therefore, on the surface, using Poisson to find special cause failures that do not conform to the requirements of Poisson use may appear to be inconsistent. But here the Poisson is used in reverse to identify data that appear not to conform to Poisson distribution requirements.

The Poisson probability distribution of events is:

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

Where:

e: An approximately 2.71828 constant, the base of the natural logarithm system.

μ: The mean number of events expected

x: A specific number of events

P(x; μ): The Poisson probability that x events are experienced, given the mean number expected is μ.

The general Poisson expression is now adapted specifically to failure events, the mean of which comes from equipment failure dates and the resulting TBF values.

$$\mu = t/MTBF \qquad (2)$$

Where:

MTBF = mean-time-between-failure

t = the specified time period, a TBF value of interest or sum of one or more consecutive TBF values.

So the Poisson distribution for failure events gives the probability of any specific number of failures x and is dependent on the time interval and MTBF, as below:

$$P(x; t/MTBF) = (e^{-t/MTBF})(t/MTBF)^{x}/x! \qquad (3)$$

3 PROBABILITY MAP

Using Poisson probability for failure time analysis involves numerous related equations that quickly become unmanageable without a systematic approach. With spreadsheet tables (probability maps) populated with appropriate equations, the entered input TBF data generates all probability values instantly. Other cells are positioned for Monte Carlo simulation and Poisson probability value (p-value) distributions for confidence intervals with a single simulation. Each event in the probability map is analyzed using only then current history and not future events. Figure 1 is the upper section of a probability map for a process pump with 13 failures. The TBF data are in column C. Column D is the running average of Column C TBF values. The p-values are given a nomenclature described by example. The probability p-v1 at, for example, cell F13 is the 10^{th} failure with a TBF value of 8 and MTBF of 466. This is a one event lookback. The equation in cell F13 is the probability of one or more events occurring within the interval. This is efficiently expressed as the complement of zero events with Excel nomenclature. So event 10 p-v1 is:

p-v1 =1-(EXP(-C13/D13)*POWER(C13/D13,0))/FACT(0)

Cell I16 is the p-v4 probability at event 13. It looks back at the TBF values for the last 4 failures. It is the probability of four or more failures within a period of 8 + 38 + 46 + 11 when the mean is 365, the dataset mean at the 13^{th} failure. So cell I16 is the probability of 4 or more events in the time of the 4 events. This is efficiently expressed as the complement of the cumulative probability of 3 events. Taking advantage of Excel’s embedded statistical formulas, cell I16 is event 13 p-v4, or:

p-v4 =1-POISSON.DIST(3,SUM(C13:C16)/D16,TRUE)

P-v4 is alarmed with 0.00021 probability. This is very strong evidence for rejecting the null hypothesis of no step change, but there is sufficient earlier evidence that could have been acted upon. The low p-v1 value in cell F13 is visually alarmed by a low p-value. This suggests the TBF value of 8 is inconsistent with a HPP with the demonstrated MTBF of 466. The Poisson interval of 8 with 1 event will occur in an HPP by random chance with only 0.01704 probability. The HPP null hypothesis can be rejected. We can say there is a step change in reliability at failure 13 with small chance of being incorrect

As happens nearly all the time in the absence of contemporaneous analysis, the step change in reliability was detected much too late – not at failure 10 but at failure 13. The opportunity to avoid the last three failures was missed. (Root cause analysis of why the reliability degradation was not detected sooner led to this current methodology).

Column / Row	B	C	D	E	F	G	H	I	J	K
3	Event	TBE	MTBE		p-v1	p-v2	p-v3	p-v4	p-v5	p-v6
4	1	13	13		0.63212					
5	2	99	56		0.82930	0.59399				
6	3	885	332		0.93026	0.79493	0.57681			
7	4	759	439		0.82253	0.88784	0.75753	0.56653		
8	5	60	363		0.15227	0.65863	0.84686	0.72991	0.55951	
9	6	503	387		0.72786	0.42757	0.66419	0.82100	0.71042	0.55432
10	7	761	440		0.82263	0.78102	0.57884	0.69565	0.80250	0.69550
11	8	1308	549		0.90788	0.89022	0.84661	0.70555	0.73861	0.78934
12	9	259	516		0.39445	0.80600	0.82740	0.79635	0.65771	0.70803
13	10	8	466		0.01704	0.11328	0.65706	0.73753	0.72795	0.59017
14	11	38	427		0.08522	0.00541	0.03598	0.52257	0.65244	0.66533
15	12	46	395		0.10995	0.01966	0.00177	0.01291	0.41035	0.57464
16	13	11	365		0.02966	0.01097	0.00241	0.00021	0.00352	0.30916

Figure 1

Process pump probability map – upper section. Event (Failure) numbers 10 thru 13 are special cause failures that do not fit the HPP null hypothesis. Strength of evidence increases to last event with p-v4 = 0.00021 probability. Early intervention could have prevented the last 3 failures

With figure 1 and subsequent similar figures that describe the three probability map sections, it is expected the reader can construct probability maps of any size for their personal use.

4 BOEING 737 MAX APPLICATION

The 737 MAX can be treated as a conceptual single system consisting of any and all reliability impacts (e.g., design, construction, testing, training, documentation, technical support). This system is equivalent to a conceptual single plane that is an in-service repairable system with number of departures between events being the Poisson interval. So the more common time between failures (TBF) is replaced with departures between events (DBE). Departures between events are estimated in this paper. Boeing and the FAA should do an analysis with exact data or best possible estimates. This paper demonstrates the methodology with estimates that are intended and believed to be conservative. By conservative is meant data that tend to understate the findings and conclusions of the paper.

4.1 Departures Between Events Estimate

The numbers and dates of 737 MAX planes delivered were periodically announced by Boeing. The basis for this analysis is the first plane was delivered on 5/16/2017 and 376 delivered by 3/10/2019 with delivery by month estimated as shown in figures 2 and 3. Announced deliveries are marked in yellow in the third column of figure 2. Production ramp up rates are generally exponential, but here we assumed the ramp up rates to be linear between the reported numbers in yellow. This is expected to yield conservative results. It was reported by a Boeing executive that by 5/21/2018, 130 planes had been delivered with 41,797 departures accumulated. This is substantially under the departure numbers generated by the estimation method, so the estimated departures appear to be conservative. It is assumed service began immediately upon delivery with the number of departures per month equal to the worldwide fleet departure rate of 94.4, calculated from Boeing published data [1].

Month	Deliveries within Month	Cummulative Deliveries	Departures within Month	Cummulative Departures
5/16/2017 1st delivery		linear ramp up	94.4/mo with immediate service	
Jun-17	10.57	10.57	997.81	997.81
Jul-17	10.57	21.14	1995.62	2993.42
Aug-17	10.57	31.71	2993.42	5986.85
Sep-17	10.57	42.28	3991.23	9978.08
Oct-17	10.57	52.85	4989.04	14967.12
Nov-17	10.57	63.42	5986.85	20953.97
Dec-17	10.57	74	6984.66	27938.62
Jan-18	11.20	85.19	8041.94	35980.56
Feb-18	11.20	96.39	9099.22	45079.78
Mar-18	11.20	107.59	10156.50	55236.27
Apr-18	11.20	118.79	11213.78	66450.05
May-18	11.20	130	12271.06	78721.10
Jun-18	28.57	158.56	14968.06	93689.17
Jul-18	28.57	187.13	17665.07	111354.24
Aug-18	28.57	215.70	20362.08	131716.32
Sep-18	28.57	244.27	23059.09	154775.41
Oct-18	28.57	272.84	25756.10	180532
Nov-18	28.57	301.41	28453.10	208984.61
Dec-18	28.57	330	31150.11	240134.72
Jan-19	19.70	349.68	33009.79	273144.51
Feb-19	19.70	369.38	34869.47	308013.98
Mar-19 (to 3/10)	6.60	376	11830.84	319845

Figure 2

Table of plane delivery estimates by month and accumulated departures by month. The cumulative departures in the last column on the right marked in yellow are those used in the analysis. See figure 3.

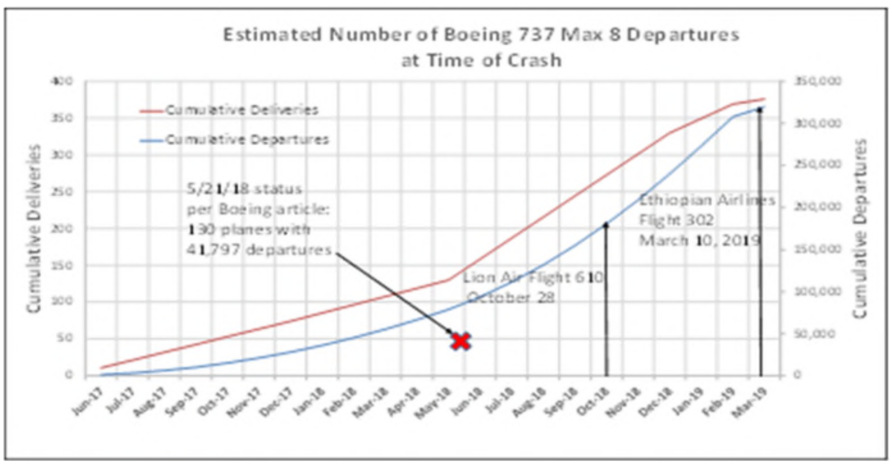

Figure 3

Graph of figure 2 data. On 5/21/18 it was reported by a Boeing executive that 41,797 departures had accumulated with 130 planes delivered. This is substantially under the departure numbers generated by the estimation method, so the estimated departures appear to be conservative.

4.2 Worldwide fleet fatal accident rate

In the process pump example, the MTBF was developed from that particular pump's history. But if the equipment has insufficient history, then MTBF of similar equipment can be used. Boeing reports 35 fatal accidents in 213.7 million departures of similar aircraft and service in a recent 10 year period [1]. There are fatal accidents that do not involve crashes. The data was a little vague so that conservatism was left in the analysis and this industry average is used. Our null hypothesis is that the 737 MAX reliability as measured by mean departures between events (MDBE) is equal to the industry mean of 6,105,714 with events generated by a process that is an HPP. This null hypothesis is a benchmark against which we compare the 737 MAX using Poisson p-values as rejection criterion. Computer simulation is used as an independent concept whose matching p-values provide confirmation of method.

4.3 Analysis and Results

Again this analysis with specific results is intended to show the method. Final results and the method used to get the results should be obtained and reported by Boeing, the FAA, or accident investigators. The basic results are presented in the upper region of figure 4. As mentioned earlier, in all but extremely small datasets, the number of equations and their relationships require an organized calculation approach. So we use the probability map for this. This upper region duplicates figure 1, with some exceptions to use worldwide mean MDBE instead of 737 MAX history, and third event probability. The middle and lower sections of the probability map relate to confidence intervals around the calculated p-values. These were omitted from figure 1 for simplicity. Cells D4 and D5 are the worldwide fleet MDBE. Upon the first event, the p-v1 is 0.02913. So about a 3% chance that the quick event would occur by random chance if the 737 MAX had average reliability as measured by MDBE. The equations for the cell values in figure 4 are found in figure 5. Row 5 in figure 4 for the second event shows p-v1 = .02256 and p-v2 = 0.00133. To refresh, p-v2 is the probability of two or more events occurring in a departures interval equal to the sum of the last two DBE. So the p-v2 value suggests there is about 1 in a thousand chance of a plane system with worldwide fleet reliability having two fatal accident events within 319,845 departures. The null is definitely rejected. The 737 MAX is statistically below worldwide mean in reliability, using data known contemporaneously with the event.

The related p-value equations in the upper region of figure 5 are based on equation 3 but using the complement, and usually in cumulative form, to simplify.

Row 6 in figure 4 is to evaluate the probability of a third event. The low p-values for the first two events cause us to reject our null hypothesis and say the 737 MAX reliability is statistically less than industry average. Cell C6 is the number of departures to be expected by the method discussed earlier for a time of 3 days. The number of planes in service for 3 days following the second event is not known, so the C6 value is based on all flying as normal. Presumably the decision by Boeing and the FAA not to ground was based on all flying. Also, the decision not to ground could have been extended beyond 3 days. The D6 value is the 737 MAX MDBE, the average of cells C4 and C5. The third event never happened but the probability of an event within 3 days is 0.04 from the probability distribution found in figure 5 cell I392 (to be discussed later).

Computer simulation is required for p-value confidence intervals and as the second confirming concept. The middle region of the probability map, rows 390 thru 392 are for treating MTBF (or MDBE) as a random variable. A random variable sample of failure time is determined by equation 4. This equation solved for t is equation 5. A random number from 0 to 1 in equation 5 returns a sample TBF or DBE.

$$F(t) = 1-e^{-t/MTBF} = RN \qquad (4)$$

RN = Random Number

$$t = -MTBF*(\ln(RN)) \qquad (5)$$

Equation 5 is in columns F and G of the probability map middle section with the samples returning a random sample MTBF or MDBE in column E. The distributions for the p-values are then found in the lower region of the probability map, rows 746 thru 748. This lower region has formulas related to a commercially available Microsoft Excel risk analysis add-in - Palisade @Risk (free trial downloads). Any other can be substituted or the term "RiskOut()+" in the equations can be deleted. Then recalculation with the F9 key will allow variability to be visually observed.

Equation 5 is also used to obtain p-values using simulation. The column at the extreme right of the probability map is added. This sums samples of consecutive TBF or DBE values. For the first event this sum on row 390 is equal to cell F390. This cell, cell F390, contains the distribution of DBEs seen in figure 6. It shows that 2.9% of the random samples from our null hypothesis is below the number of departures at the time of the first event. This matches the Poisson probability. While Poisson and computer simulation are two completely different concepts giving equal answers, they are not independent. Equation 3 and 5 are related.

Col / Row	B	C	D	E	F	G	H	DBE Random Samples
	Event	TBE	MTBE		p-v1	p-v2		
4	1	180,532	6,105,714		0.02913			
5	2	139,313	6,105,714		0.02256	0.00133		
6	3	3,549	159,923		0.02195			
	Event	TBE	MTBE	MTBE RV	TBE-1	TBE-2		
390	1	180,532	6,105,714	5,083,569	5,083,569			5,083,569
391	2	139,313	6,105,714	16,891,216	31,135,026	2,647,405		33,782,431
392	3	3,549	159,923	7,284	7,284			7,284
	Event	TBE	MTBE	MTBE RV	p-v1	p-v2		

Figure 4

Probability map with equations in cells shaded blue modified or deleted for the 737 MAX data. The column on the right is added for simulation, a second concept to check Poisson values. Related cell equations are found in Figure 5.

Poissson Probability Map - modified for Boeing 737 MAX analysis							Additional Simulation
Col / Row	B	C	D	E	F	G	I
	Event	DBE	MDBE		p-v1	p-v2	
4	1	180532	6105714		=1-(EXP(-C4/D4)*POWER(C4/D4,0))/FACT(0)		
5	2	139313	6105714		=1-(EXP(-C5/D5)*POWER(C5/D5,0))/FACT(0)	=1-POISSON.DIST(1,SUM(C4:C5)/D5,TRUE)	
6	3	3549	=AVERAGE(		=1-(EXP(-C6/D6)*POWER(C6/D6,0))/FACT(0)		
Poissson Probability Map - modified for Boeing 737 MAX analysis							
	Event	DBE	MDBE	MDBE RV	DBE-1 RV	DBE-2 RV	
390	1	=C4	=D4	=AVERAGE(F390:F390)	=RiskOutput()+-$D390*LN(RAND())		
391	2	=C5	=D5	=RiskOutput()+AVERAGE(F391:G391)	=RiskOutput()+-$D391*LN(RAND())	=-$D391*LN(RAND())	=RiskOutput()+SUM(F391+G391)
392	3	=C6	=D6	=RiskOutput()+AVERAGE(F392:G392)	=RiskOutput()+-$D392*LN(RAND())	=-$D392*LN(RAND())	=RiskOutput()+-E392*LN(RAND())
	Event	DBE	MDBE	MDBE RV	p-v1 RV	p-v2 RV	
746	1	=C4	=D390	=E390	=RiskOutput()+1-(EXP(-C746/E746)*POWER(C74		
747	2	=C5	=D391	=E391	=RiskOutput()+1-(EXP(-C747/E747)*POWER(C74	=RiskOutput()+1-POISSON.DIST(1,SUM(C746:C747)/E747,TRUE)	
748	3	=C6	=D392	=E392	=RiskOutput()+1-(EXP(-C748/E748)*POWER(C74		

Figure 5

Equations for figure 4 probability map. Some cells are modified or deleted to accommodate 737 MAX application.

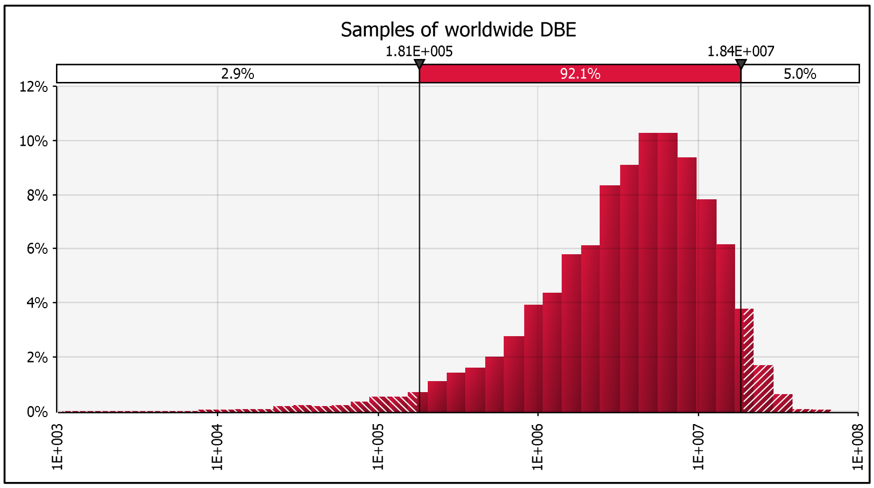

Figure 6

Computer generated DBE samples from the null hypothesis. 2.9% are below the number of departures at the time of the first event. This suggest the first event departures to failure is relatively rare by random chance. It matches the p-v1 value in figure 4 cell F4.

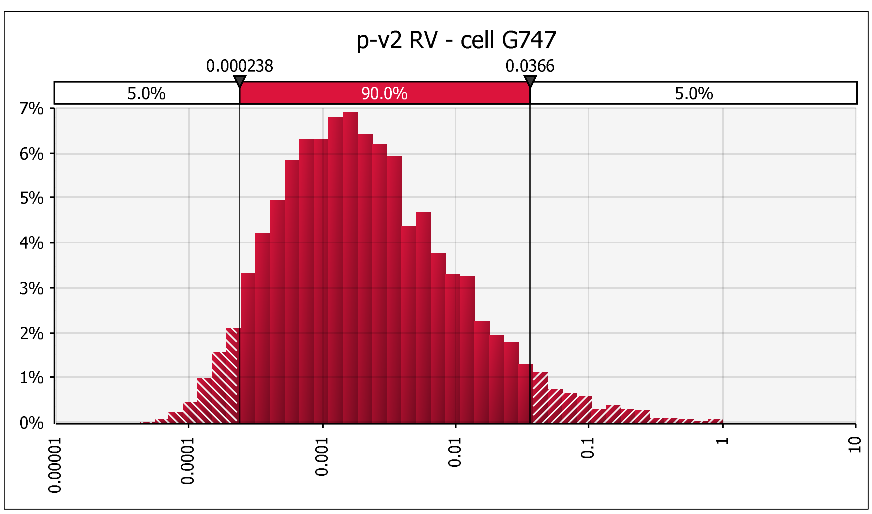

Figure 7

Probability value distribution for 2 events within 319,845 departures. Any confidence interval is available from the distribution of p-values.

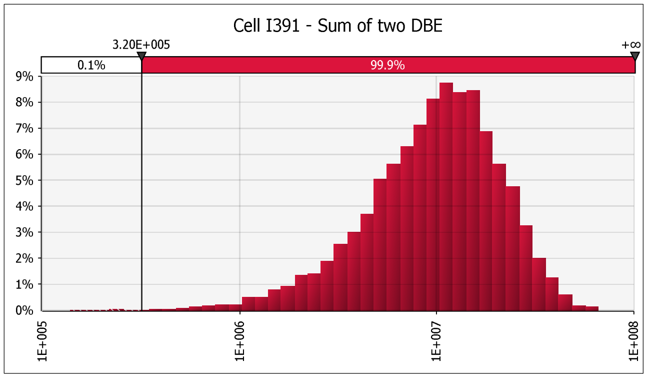

Figure 8

The worldwide fleet probability of 2 consecutive events occurring within the number of departures experienced at the time of event 2 is 0.1%, matching the probability value p-v2 in figure 4 cell G5.

At event 2, the p-value of most interest is the very low p-v2 value. The probability distribution for p-v2 is in cell G747 and is seen in figure 7. The 90% confidence interval is 0.000238 to 0.0366. This is checked with simulation and shown in figure 8 to be 0.1%. The simulation p-value is consistent with the Poisson p-value confidence interval.

The null hypothesis that the 737 MAX is equal to industry average is rejected. Upon the first event, failure time analysis would ideally have been done and the data considered, communicated and discussed. Upon the second event, failure time analysis would clearly have led to immediate grounding. It is expected that Boeing, the FAA or accident investigator analysis with exact input data will lower all p-values, thus increasing the gap between the 737 MAX and the worldwide fleet.

Now we consider the risk of a third event during the three days of operation before grounding. Cell F6 does not reflect that we only have two data points forming our mean. Equation 5 is used twice to get the DBE distribution in cell I392. In 10,000 iterations, 4% were below 3,549, the number of departures expected in 3 days with all 737 MAX in service (see figure 9). The probability of a third event in 3 days is 4%.

This high risk of a third event is taken as proof that Boeing and the FAA did not conduct this or similar analysis. Failure time analysis is typically not used for in-service repairable systems (and the author supposes likewise for any other HPP in business, medical, etc.). When it is conducted, it is too late to be of greatest value. Both of these tendencies have created a restraining paradigm regarding failure time analysis that should be corrected. The reliability, technical and data analysis communities are challenged to do so.

ACKNOWLEDGEMENTS

Appreciation to prior co-authors client reliability engineers Landon Davis and Oliver Sac and Zachry Senior Director of Operations Kevin Bordelon. Also client reliability manager Darrell Evans and reliability engineer Vu Phan for their progressive support for automated failure time analysis.

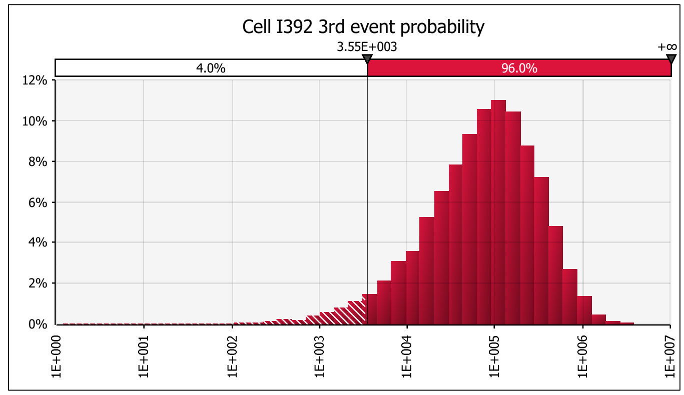

Figure 9

Probability of a 3^rd^ event in 3 days from cell I392 is 4%.

REFERENCES

1. Jan Smith, Oliver Sac, Kevin Bordelon, "Contemporaneous Failure Time Analysis Using Poisson Probability", Annual Reliability and Maintainability Symposium, 2018

2. Boeing, "Statistical Summary of Commercial Jet Airplane Accidents Worldwide Operations", 2018

BIOGRAPHY

Jan B. Smith, P.E.
Zachry Industrial, Inc.
3741 Red Bluff Rd.
Pasadena, TX 77503 USA

email: smithjb@zachrygroup.com

Jan B. Smith, P.E., BSME, Senior Reliability Consultant for Zachry Industrial, Inc. supports reliability engineering in client process industry plants. Principle work in this role includes asset reliability strategies including FMEA, and statistical and probabilistic analysis for failure rate trends and process unit availability and capacity. He has chaired conferences on plant reliability, authored several technical papers on the subject, and has developed and taught in house and public seminars on root cause analysis. He holds three patents on process unit availability and capacity assessment using empirical probabilistic methods. His 54 year career has been devoted to developing and applying reliability engineering principles and practices in the process industry, with emphasis on probabilistic use of observational data.