A Few Risk Analysis Tools

1

Hartford section ASQ, Short Course, October, 2017

Stephen N. Luko
United Technologies Aerospace Systems

FMEA

- Probably familiar to all
- Several variations: PFMEA, DFMEA, FMECA
- Fundamentals are similar in each
- Fundamentals: failure mode, Frequency, Severity, Detectability; risk prioritization number (RPN); RPN=FXSXD; smallest is best'; Pareto of RPN/effects
- Can take a long time to complete in complex products/processes
- Now a standard in many industries

FMEA; Frequency Scale

RATING	PROBABILITY OF FAILURE	POSSIBLE FAILURE RATES	ASSOCIATED PROCESS Cp's
1	Remote: Failure is unlikely.	< 1 in 10^6	Cp > 1.33
2	Low: Relatively few failures.	< 1 in 20,000	Cp > 1.33
3	Low: Relatively few failures.	< 1 in 4000	Cp > 1.00
4	Low: Relatively few failures.	< 1 in 1000	Cp > 1.00
5	Moderate: Occasional failures.	< 1 in 400	Cp < 1.00
6	Moderate: Occasional failures.	< 1 in 80	Cp < 1.00
7	Moderate: Occasional failures.	< 1 in 40	Cp < 1.00
8	High: Repeated failures.	< 1 in 20	Cp < 1.00
9	High: Repeated failures.	< 1 in 8	Cp < 1.00
10	Very High: Failure is almost inevitable.	> 1 in 8	Cp < 1.00

FMEA; Severity Scale

Category (Product)	Criteria: Severity of Effect (Effect on Product) – DFMEA & PFMEA	Rank	Category (Process)	Criteria: Severity of Effect (Effect on Process) - PFMEA
Safety and/or Regulatory Compliance	Potential failure mode affects safe vehicle operation and/or involves noncompliance with government regulation without warning.	10	**Safety and/or Regulatory Compliance**	May endanger operator (machine or assembly) without warning.
	Potential failure mode affects safe vehicle operation and/or involves noncompliance with government regulation with warning.	9		May endanger operator (machine or assembly) with warning.
Primary Function *Essential*	Loss of primary function (vehicle inoperable, does not affect safe vehicle operation)	8	**Major Disruption**	100% of product may have to be scrapped. Line shutdown or stop ship.
	Degradation of primary function (vehicle operable, but at reduced level of performance)	7	**Significant Disruption**	A portion of the production run may have to be scrapped. Deviation from primary process; decreased line speed or added manpower.
Secondary Function *Convenient*	Loss of secondary function (vehicle operable, but comfort / convenience functions inoperable)	6	**Rework out-of-station**	100% of production run may have to be reworked off line and accepted.
	Degradation of secondary function (vehicle operable, but comfort / convenience functions at reduced level of performance)	5		A portion of the production run may have to be reworked off line and accepted.
Annoyance	Appearance or Audible Noise, vehicle operable, item does not conform. Defect noticed by most customers (> 75%)	4	**Rework in-station**	100% of production run may have to be reworked in station before it is processed.
	Appearance or Audible Noise, vehicle operable, item does not conform. Defect noticed by many customers (50%)	3		A portion of the production run may have to be reworked in-station before it is processed.
	Appearance or Audible Noise, vehicle operable, item does not conform. Defect noticed by discriminating customers (< 25%)	2	**Minor Disruption**	Slight inconvenience to process, operation, or operator
No effect	No discernible effect.	1	**No effect**	No discernible effect

FMEA; Detectability Scale

RATING	PROBABILITY OF DETECTION	LIKELIHOOD OF DETECTION
1	Very high	Will almost certainly detect a potential design weakness prior to release for production or process. Process controls (i.e., final test, SPC, etc.) will almost certainly detect the existence of a defect (process automatically detects failure).
2 3	High	Good chance of detecting a potential design weakness prior to release for production. Will be detectable after release but before mass production. Process controls have a good chance of detecting the existence of a defect.
4 5 6	Moderate	May detect a potential design weakness prior to release for production. Will be detectable prior to first shipment. Process controls may detect the existence of a defect.
7 8	Low	Not likely to detect a potential design weakness. Failure mode will be detected prior to occurring in the field. Process controls have a poor chance of detecting the existence of a defect.
9 10	Very low	Probably will not detect a potential design weakness until field failures occur. Process controls probably will not detect the existence of a defect.

Redundancy - 1

- The "***principle of redundancy***"
- Redundancy can reduce the risk of a failure in classical engineering design; in inspection effectivity; in safety checks
- POD (probability of detection); Numerous interpretations of this: multiple checking; etc.
- Even where the risk probability changes from time to time, redundancy will reduce risk.

Redundancy - 2

- In inspections: Risk probability of failing to detect a failure is p; in n independent repetitions the risk probability reduces to p^n. For p=0.1 and n=3, the risk probability is 0.001.
- Numerous interpretations of this: POD (probability of detection); multiple checking; etc.
- Even where the risk probability changes from time to time, redundancy will reduce risk.

Redundancy - 3

- ***Reverse redundancy*** (my term) occurs when multiple tasks occur simultaneously. There is a higher risk (probability) of a mistake when "multitasking" occurs than when one task is done.
- Normalized "event probability". If the "event" probability is changing from time to time in some repeated task, what is the equivalent normalized or "average" event probability over the several trials? Example.

$$p = \sqrt[n]{p_1 p_2 p_3 \ldots p_n} \qquad p = \sqrt[5]{(0.2)(0.5)(.8)(0.05)(0.1)} = 0.21$$

Checklists - 1

- Probably the simplest but certainly very effective means of reducing risk in any type of process.
- Particularly effective when redundancy is used (executing a checklist multiple times or using a series of checklists)
- Uses: process preparation or setup, safety checks, auditing checklist, workday/job checklist;
- Quality – SIPOC – use in conjunction with.

Checklists - 2

Example:

CheckList Master for problem solving	
Tools and other items to consider	
item	
1	What problem are we trying to solve? What is the Background?
2	What kind of data do we have?
3	How much data do we have?
4	What is the "quality" of the data?
5	How was the data obtained (sampling/collecting method)?
6	What was the "measurement system" used?
7	How is the data organized?
8	Who manages/owns the data?
9	Will there be more data forthcoming?
10	Is there any missing data?
11	Does this data "represent" the problem we are faced with? (see 5)
12	What analysis tools will we use on this data?
13	What summarization displays will we use for reporting ?

Ring Fence (redundancy) - 1

- A form of redundancy in sampling inspection that incorporates quantitative detection probability at the various stages of inspection prior to reaching a customer's hands.
- Two short papers for the presenter.


DATA POINTS
A STATISTICS Q&A

Inspection Redundancy – Part 1

Why multiple or redundant inspections make sense.

By Stephen N. Luko


DATA POINTS
A STATISTICS Q&A

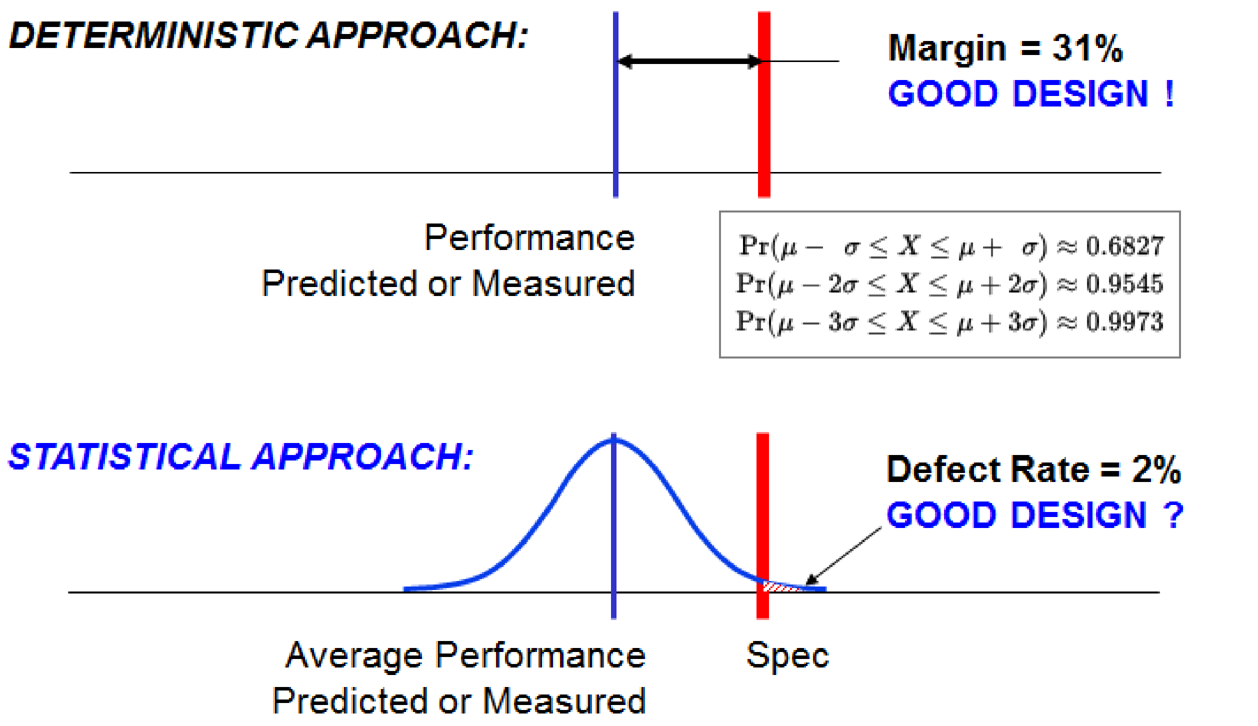
Inspection Redundancy – Part 2

Multiple inspections at different locations or at different points are a consideration with inspection redundancy.

By Stephen N. Luko

Design for Variation (DFV) - 1

Two Approaches to Design:



Design for Variation (DFV) - 2

How to Decrease Variation

- Improve the Existing Process via *Process Certification*
- Replace the Existing Process with New Capability:
  – Different Supplier
  – New Manufacturing Technology
- A Mix of Both

Cost – Benefit Analysis Required

Forecasting - 1

- **What happens next?** What we expect to happen based on an existing “fleet” of fielded units; a statistical model (based on current field data); and a utilization going forward.
- We can do this by month/week/year; or by some amount of future time.
- There are many variations of the technique

Forecasting - 2

- If F(t) is the cumulative distribution function of the model being used; t the time on a unit in the field and Δt a future time, then the forecasted failure probability in the next interval of time (Δt) is:

$$r = \frac{F(t+\Delta t) - F(t)}{1 - F(t)}$$

- Quantity "r" is the probability of item failure by time t+ Δt, given survival at time t. When a "fleet" of items is involved we simply add up this quantity over the "fleet". That number is the expected future failures.

Forecasting - 3

Example: Weibull Model; 76 units in fleet; variable times; utilized at 100 hours (future time)

	parameters					
beta	1.500	*enter*		GO TO Catalog	Clear Data	
charLife	1,500.0	*enter*				
delta (period)	100		*enter*			
total Risk	3.053	*calculated*		*total units*	**76**	*calculated*
sigma	1.747	*calculated*		*total time period*	**7,600**	*calculated*
starting, n	**76**	*calculated*		*total time, adjusted*	**7,295**	*calculated*
remaining	**72.95**	*calculated*		*AST*	**239**	*calculated*
med remaining t	**1006.211**	*calculated*		*rate*	**4.186E-04**	*calculated*
F(AST+delta)	**0.10207**	*calculated*		*AST=average suspension time*		

		end time				
t	**Frequency**	**t+delta**	**F(t)**	**F(t+delta)**	**unit risk**	**Freq Risk**
100	12	200.0	0.01706596	0.04752027	0.03098306	0.37179678
500	15	600.0	0.17506451	0.22351831	0.05873647	0.88104702
300	25	400.0	0.08555936	0.12864523	0.04711719	1.17792982
250	8	350.0	0.06577819	0.10659092	0.04368634	0.34949074
0.00	16	100.0	0.00000000	0.01706596	0.01706596	0.27305532

Forecasting - 4

Example (continued): Weibull Model; 76 units in fleet; variable times; utilized at 20 hours/mth for the next year

Month	utilization	start	end	risk	CumeRisk	Cume Rate
1	20	0.0	20.0	0.555914	0.555914	3.6573E-04
2	20	20.0	40.0	0.589728	1.145642	3.7686E-04
3	20	40.0	60.0	0.615102	1.760744	3.8613E-04
4	20	60.0	80.0	0.636746	2.397491	3.9432E-04
5	20	80.0	100.0	0.655829	3.053320	4.0175E-04
6	20	100.0	120.0	0.672908	3.726228	4.0858E-04
7	20	120.0	140.0	0.688313	4.414541	4.1490E-04
8	20	140.0	160.0	0.702263	5.116804	4.2079E-04
9	20	160.0	180.0	0.714918	5.831722	4.2630E-04
10	20	180.0	200.0	0.726399	6.558121	4.3146E-04
11	20	200.0	220.0	0.736804	7.294925	4.3630E-04
12	20	220.0	240.0	0.746211	8.041136	4.4085E-04

Statistical Intervals

- **Statistical Intervals** provide information on the uncertainty of reported results as "point" estimates. There are several scenarios which these intervals can apply to.
- Many people are familiar with, or have heard of the so called "**confidence interval**"; but the other two types of intervals, the **prediction and tolerance type intervals** are not well known.

Confidence type intervals - 1

- **A Confidence Interval** applies to a parameter of a model. For example, the mean or the standard deviation or the 5th percentile of a normal distribution, the proportion non conforming in a process output, the rate of occurrence of a certain defect type. Many other examples can be given.
- When a parameter is estimated it is usually possible to calculate a standard error for the estimate. This is traditionally called "***the standard error of the estimate***" or **SE**.

Confidence type intervals - 2

- The SE is interpreted as the estimate of the standard deviation of a large set of estimates of the same type of statistic in repeating the sampling process over and over again, under the same conditions, and using differing data each time. It is also assumed that the distribution of these estimates is asymptotically normally distributed. From these facts, the confidence interval for the parameter is constructed as:

$$\hat{\theta} \pm Z_{\alpha/2} SE(\hat{\theta})$$

Confidence type intervals - 3

$$\hat{\theta} \pm Z_{\alpha/2} SE(\hat{\theta})$$

- The unknown parameter is θ; the "hat" notation over the parameter means "an estimate of"; and "SE" is the estimated standard error. The two sided interval (in this case) is said to carry a confidence level of 100(1-α)%. The quantity $Z_{\alpha/2}$ is a positive value from the standard normal distribution for which $P(-Z_{\alpha/2} \leq Z \leq Z_{\alpha/2})=1-\alpha$.

Confidence type intervals - 4

$$\hat{\theta} \pm Z_{\alpha/2} SE(\hat{\theta})$$

- The Interval so constructed would contain the value of the unknown parameter in the long run, a proportion of the time, 1-α, if made under the same conditions and with differing data, many times.
- **Example:** For an unknown mean, μ, the interval is:

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

- “t” is student’s t; “x-bar” and s are the sample mean and standard deviation.

Prediction type intervals - 1

- The **Prediction Interval** applies to a future observation such as a future individual observation or a future sample mean. They can also be applied to a future set of k observations (where k>1).
- Prediction intervals and be parametric (for example, the normal distribution is used) or non-parametric (applies to any distribution). They also apply to variable and discrete type data (for example binomial and Poisson count data).

Prediction type intervals - 2

- To use a **two sided Prediction Interval method, where the normal distribution is being used:**
- Specify a confidence, C=1-α, and k (the number of future observations). When σ is unknown (most of the time) use the sample standard deviation, s.

$$\bar{x} \pm t_{\alpha/(2k)} s \sqrt{1 + \frac{1}{n}}$$

- $t_{\alpha/(2k)}$ is the positive one student's t with n-1 degrees of freedom and n is the initial sample size. If σ is known, replace t with Z (standard normal).

Prediction type intervals - 3

ByGroup	xbar	s	n	LL(2.5,1)	UL(97.5,1)	LL(2.5,5)	UL(97.5,5)
1	101.39	10.38	10	76.75	126.02	65.99	136.78
2	104.15	12.85	10	73.66	134.64	60.34	147.96
3	101.39	10.85	10	75.65	127.13	64.41	138.38
4	101.71	10.26	10	77.37	126.06	66.73	136.69
5	95.29	6.52	10	79.83	110.75	73.07	117.50
6	93.20	8.05	10	74.10	112.31	65.75	120.66
7	95.89	8.15	10	76.56	115.23	68.11	123.67
8	107.77	9.14	10	86.10	129.44	76.63	138.91
9	101.38	10.62	10	76.19	126.57	65.19	137.58
10	92.71	8.89	10	71.61	113.81	62.40	123.03

Ten sets of n=10 observations (groups); with calculated prediction limits (95%) for the next 1 and the next 5 observations)

x1	x2	x3	x4	x5	one	five
91.97	112.02	95.46	94.81	95.45	1	1
101.89	96.95	96.98	92.34	98.94	1	1
90.68	102.58	79.07	104.79	101.39	1	1
119.25	104.00	111.54	103.03	88.36	1	1
92.69	105.35	85.85	100.56	115.20	1	1
100.22	96.25	104.63	96.40	80.05	1	1
88.34	97.59	102.33	79.87	121.35	1	1
89.15	86.46	104.33	84.69	99.94	1	1
92.69	121.05	103.84	83.90	87.71	1	1
104.23	105.67	83.24	106.47	99.27	1	1

Here are the next 5 observations. In 100 cases, 98% are correct for the next single observation and 96% correct for the next 5 observations. Remember that we are using 95% confidence).

% true, 1	% true, 5
98	96

Prediction type intervals - 4

- The non-parametric method case can require much more data for a desired interval. There are also many variations on this theme. We use the min and max of the initial sample for this illustration.
- For an initial sample size , n, the confidence that the next k observations fall between the sample initial sample min and max is:

$$C = \frac{n(n-1)}{(n+k)(n+k-1)}$$

Prediction type intervals - 5

- Example: the non-parametric case. Suppose n=9 initial observations. We want to use the sample min and max as the future interval for the next 3 observations. With what confidence?

$$C = \frac{9(9-1)}{(9+3)(9+3-1)} = 0.54$$

- Thus, 54% confidence. Suppose we wanted 90% confidence, continuing to use k=3. Use the formula and solving for the smallest n, find n=58 works.

Prediction type intervals - 6

- Of course if the sample size is larger, confidence will be larger sooner. Use n=100, k=3

$$C = \frac{100(100-1)}{(100+3)(100+3-1)} = 0.942$$

- Nearly 95% !

Tolerance type intervals - 1

- The **Tolerance Interval** applies to an entire population or process. Again, we can have parametric and non-parametric cases for variable and discrete type data.
- A tolerance intervals needs two specified values: a) confidence; and b) the desired proportion, p, to be captured in the population. In the case of the normal distribution we just have to look one number, k, in a table as a function of C and p, and the sample size n.

Tolerance type intervals - 2

- The normal distribution case with estimated mean and standard deviation and sample size n. A short table for n=30 is shown below.

n	C	p	k
30	0.90	0.900	2.0252
30	0.90	0.950	2.4132
30	0.90	0.990	3.1714
30	0.90	0.997	3.6540
30	0.90	0.999	4.0514
30	0.95	0.950	2.5496
30	0.95	0.990	3.3508
30	0.95	0.997	3.8606
30	0.95	0.999	4.2805
30	0.99	0.990	3.7345
30	0.99	0.997	4.3027
30	0.99	0.999	4.7707

The following interval is a tolerance interval for 100p% of the entire population/process at confidence 100C%.

$$\bar{x} \pm ks$$

Notice that people often use "3 sigma" and claim 99.7% captured for the normal case. But this chart shows that when n=30, we should really use 3.86 sigma if we wanted to claim this with say 95% confidence.

Tolerance type intervals - 3

- For the case of a two sided interval bounded by the min and max in a sample, at least 100p% of the population lies in the interval with confidence C, when a sample size of n is used. Analysis of this case leads to the following equation involving C, p and n.

$$np^{n-1} - (n-1)p^n \geq 1 - C$$

This equation can be solved for any parameter (n, C, p) whenever any two of them are specified.

Tolerance type intervals - 4

$$np^{n-1} - (n-1)p^n \geq 1 - C$$

n	p	C	n	p	C
9	0.600	0.9295	388	0.990	0.9004
10	0.600	0.9536	473	0.990	0.9502
14	0.600	0.9919	662	0.990	0.9901
15	0.750	0.9198	777	0.995	0.9002
18	0.750	0.9605	947	0.995	0.9500
24	0.750	0.9910	1,325	0.995	0.9900
38	0.900	0.9047	1,440	0.9973	0.9002
46	0.900	0.9520	1,756	0.9973	0.9501
64	0.900	0.9904	2,456	0.9973	0.9900
77	0.950	0.9027	3,889	0.9990	0.9000
93	0.950	0.9500	4,742	0.9990	0.9500
130	0.950	0.9900	6,636	0.9990	0.9900

The table shows the confidence achieved using a sample of size n for the claim that the largest and smallest order statistics (min and max) will bound a proportion of at least p of the population.

Tolerance type intervals - 5

$$np^{n-1} - (n-1)p^{n} \geq 1 - C$$

<u>*Example:*</u> In a sample of n=31, with what confidence coefficient may we claim that the sample range covers at least 90% of the population? In (3) use n=31 and p=0.9 and solve for C. The confidence coefficient is C = 0.830. There is thus an 83.0% confidence that the sample range will cover 90% of the parent population when using n=31.

<u>*Example:*</u> Determine the sample size necessary to claim that a population proportion p=0.95 or more will be covered by a sample range with confidence 0.90. In (3) use p=0.95 and C=0.9 and solve for n by iteration. The sample size that just achieves this is n=77.

In using this, we are assuming we are sampling from a process in statistical control or a random sample from a large lot' although we are not assuming any kind of distribution

Tolerance type intervals - 6

- For a one sided interval, things are simpler. At least 100p% of the population lies either above the min or below the max (as the case may be), with confidence C and using a sample size of n.

$$p^n \geq 1 - C$$

This equation can be solved for any parameter (n, C, p) whenever any two of them are specified.

For example: Should we want to use 95% confidence and claim that a proportion of at least p=0.99 lies above $x_{(1)}$, then using the equation we find that n=299 will just achieve this.

Tolerance type intervals - 7

$$p^n \geq 1 - C$$

C	0.9000	0.9500	0.9600	0.9700	0.9800	0.9900	0.9950	0.9990
0.500	7	14	17	23	35	69	139	693
0.550	8	16	20	27	40	80	160	799
0.600	9	18	23	31	46	92	183	916
0.632	**10**	**20**	**25**	**33**	**50**	**100**	**200**	**1,000**
0.700	12	24	30	40	60	120	241	1,204
0.750	14	28	34	46	69	138	277	1,386
0.800	16	32	40	53	80	161	322	1,609
0.850	19	37	47	63	94	189	379	1,897
0.900	**22**	45	57	76	114	230	460	2,302
0.910	23	47	59	80	120	240	481	2,407
0.920	24	50	62	83	126	252	504	2,525
0.930	26	52	66	88	132	265	531	2,658
0.940	27	55	69	93	140	280	562	2,813
0.950	29	**59**	74	99	149	299	598	2,995
0.960	31	63	**79**	106	160	321	643	3,218
0.970	34	69	86	**116**	174	349	700	3,505
0.980	38	77	96	129	**194**	390	781	3,911
0.990	44	90	113	152	228	**459**	919	4,603
0.991	45	92	116	155	234	469	940	4,709
0.992	46	95	119	159	239	481	964	4,826
0.993	48	97	122	163	246	494	990	4,960
0.994	49	100	126	168	254	510	1,021	5,114
0.995	51	104	130	174	263	528	**1,058**	5,296
0.996	53	108	136	182	274	550	1,102	5,519
0.997	56	114	143	191	288	579	1,159	5,807
0.998	59	122	153	205	308	619	1,240	6,212
0.999	66	135	170	227	342	688	1,379	**6,905**
0.9999	88	180	226	303	456	917	1,838	9,206

- For one sided intervals: variables data uses the most extreme value (min or max) as the boundary; attribute data requires zero defectives.

Tolerance type intervals - 8

Sample Size, n, Required to Contain at least p% of a Population, between the largest and smallest order statistic, With Confidence, C%										
	Confidence level %									
p%	50	75	90	95	96	97	98	99	99.5	99.9
50.0	3	5	7	8	8	8	9	11	12	14
60.0	4	6	9	10	11	12	12	14	16	19
70.0	6	9	12	14	15	16	17	20	22	27
75.0	7	10	15	18	18	20	21	24	27	33
80.0	9	13	18	22	23	25	27	31	34	42
85.0	11	18	25	30	32	34	37	42	47	58
90.0	17	27	38	46	49	52	57	64	72	89
95.0	34	53	77	93	99	105	115	130	146	181
96.0	42	67	96	117	124	132	144	164	183	227
97.0	56	89	129	157	166	177	193	219	245	304
98.0	85	134	194	236	249	266	290	330	369	458
99.0	168	269	388	473	500	534	581	661	740	919
99.5	336	538	777	947	1,001	1,047	1,165	1,325	1,483	1,842
99.7	622	997	1,440	1,756	1,855	1,982	2,559	2,456	2,749	3,414
99.9	1,678	2,692	3,889	4,742	5,011	5,354	5,832	6,635	7,426	9,224

- For one two sided intervals: variables data uses the min and max as the boundary; attribute data requires zero defectives between two given values.

Binomial Distribution - 1

- X is a number of "defective" units (or nonconforming units) in a sample of n – usually from a process or a large lot.
- There is a proportion, p, of units in the process/lot that have the attribute in question (usually unknown or is assumed).
- Sample to sample independence (identical conditions.
- Often we want $P(X \geq k)$ or $P(X \leq k)$ in a sample of n.

Binomial Distribution - 2

- Probability mass function or the probability of exactly X in the sample.

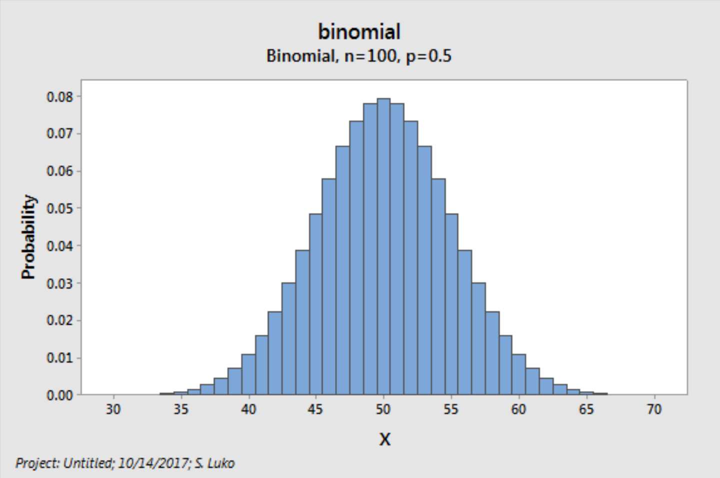
$$f(x) = \frac{n!}{(n-x)!\,x!} p^x (1-p)^{n-x}, \quad for\ x = 0,1,2,3,\ldots n$$

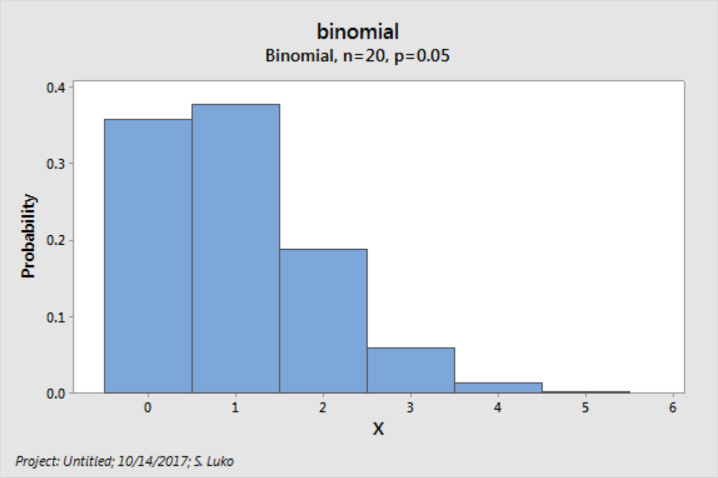
- Mean and variance: $E(x) = np, \quad var(x) = np(1-p)$

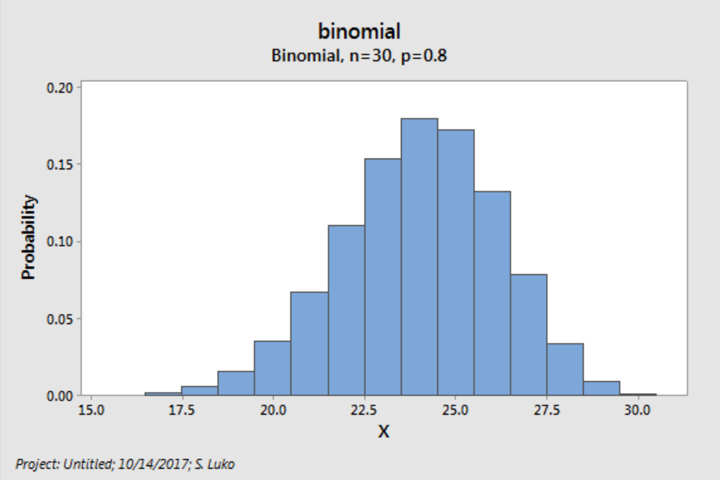
- Cumulative distribution function:

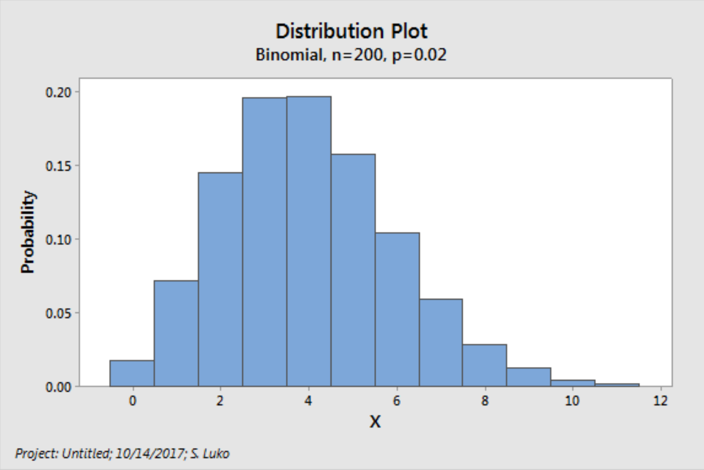
$$F(x) = P(X \leq x) = \sum_{y=0}^{x} f(y)$$

Binomial Distribution - 3









Binomial Distribution - 4

- Calculations with a binomial distributions can be done:
  - By hand with a calculator (lengthy)
  - Using a table (limited)
  - Using excel (function is built in)
  - Using Minitab (functions/graphics are built in)
  - Using a TI graphics calculator
  - Other: many programs and calculators do this.
  - Examples.

Binomial Distribution - 5

- Things that are binomially distributed
    - Tossing a coin n times; the number of heads.
    - The number of defective units in a sample of n in a stable process
    - Make 300 sales calls; the number who listen
    - n people walk into a store, x people buy something
    - Number of flights that depart on time in a certain location in a total o n flights.
    - Many Other: games, gambling, sampling/risk typ problems

Binomial Distribution - 6

- ***Example*** *– An airline knows that 5% of all passengers don't show up for their flight. As a profit strategy, it is trying to decide on selling extra tickets to recoup the recovered seats when customers don't show up. Suppose it decides to sell 52 tickets on a flight that seats 50. What is the probability that every passenger that shows up has a seat? (note: management assumes that passengers decide to show up or not, independently).*
- *Specify the problem in probabilistic terms.*
- *Is this a binomial model? In what way? What would be n and p? Given that it is binomial, what are the parameters and probability calculation are we making?*
- *What is the "risk" in this scenario?*

Binomial Distribution - 7

- ***Analysis/answers*** *– It is a binomial model. n=52 tickets are sold and p=0.95 is the probability each shows up to take a place on the Aircraft. Parameters are n=52 and p=0.95. We want the probability P(X≤50) where X is the number pf people who show up.*
- ***The risk to the airlines is: P(X≥51)***

$$P(X \geq 51) = \frac{52!}{(52-51)!51!} 0.95^{51}(1-p)^{52-51} + 0.95^{52} = 0.259$$

- ***The probability that every pass anger who shows up has a seat is: P(X≤50) which is 1-(X≥51) or 1-0.259 =0.741***

- *One can further see the business risk in that a customer who has purchased a ticket but has no seat will be very disappointed. The Airlines will have to compensate this person substantially in order to erase the disappointment and perhaps maintain customer loyalty*

Poisson Distribution - 1

- The **Poisson distribution** is probably the most important and widely used statistical model for discrete events in all of statistics.
- The random variable X is counting the number of events that occur in a fixed period of “time” t.
- “events” are of the random type and there is no limit to the number of such events that could occur in the interval t.
- A “random” event is one that is not age dependent and depends only on how long its been since the last event.

Poisson Distribution - 2

- Probability mass function or the probability of exactly X in the sample.

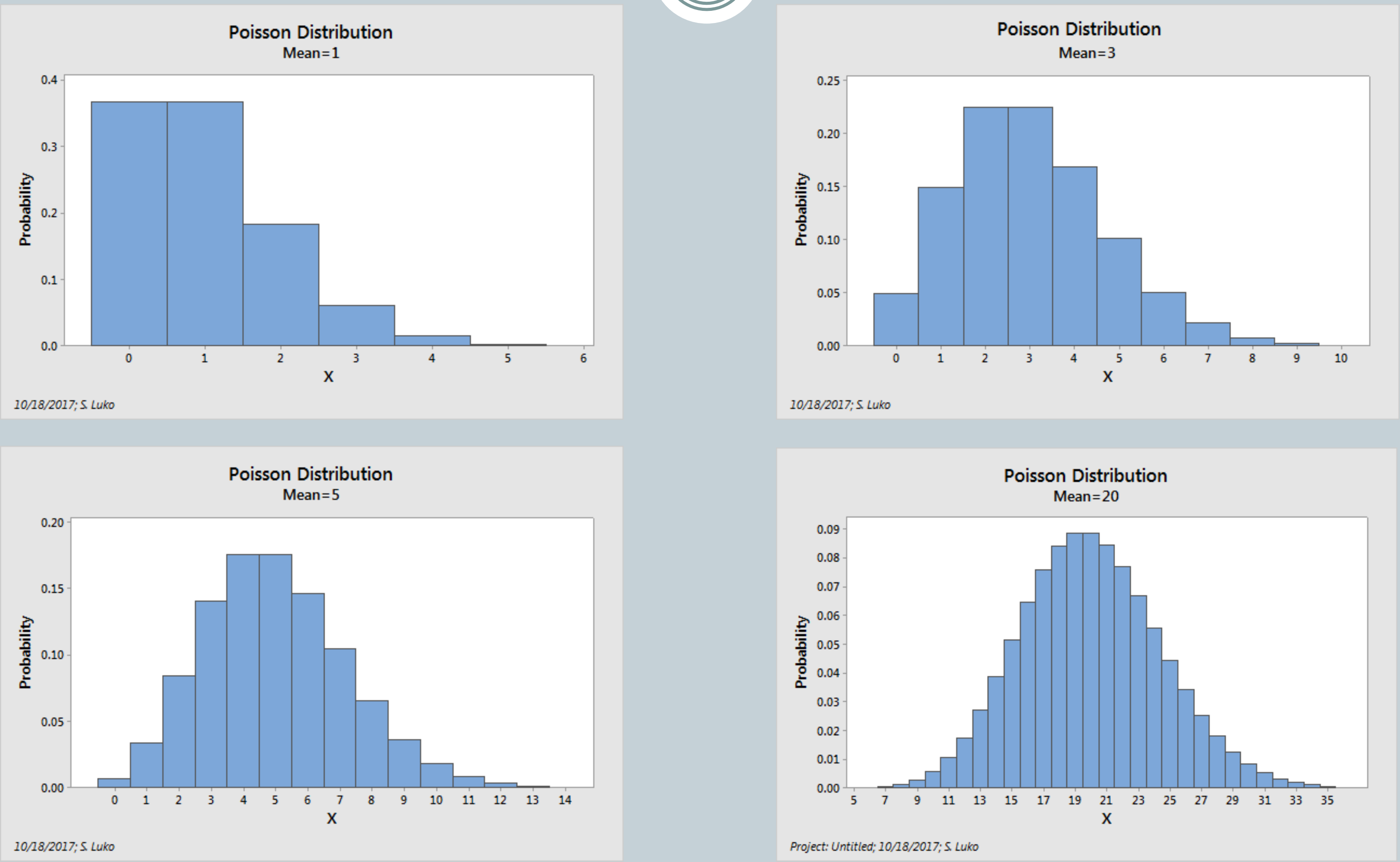
$$f(x) = \frac{e^{-\lambda t}(\lambda t)^x}{x!} \quad for\ x = 0,1,2,3,\ldots n$$

- Mean and variance: $E(x) = \lambda t, \ \ var(x) = \lambda t$

- Cumulative distribution function:

$$F(x) = P(X \le x) = \sum_{y=0}^{x} f(y)$$

Poisson Distribution - 3



Poisson Distribution - 4

- Calculations with a Poisson distributions can be done:
  - By hand with a calculator (lengthy)
  - Using a table (limited)
  - Using excel (function is built in)
  - Using Minitab (functions/graphics are built in)
  - Using a TI graphics calculator (awkward)
  - Other: many programs and calculators do this.
  - Examples.

Poisson Distribution - 5

- Things that are Poisson distributed
  - Rare events that happen randomly in an interval.
  - Approximations of a binomial when p is small and n is large (sampling).
  - Number of surface imperfections on sheet metal stock.
  - Number of defects on a bearing inner race surface.
  - Accidents (all kinds); extreme weather events
  - Count or rate of Employee attrition.

Poisson Distribution - 6

- ***Example*** – *In a large facility with many people, the number of lost time accidents for all reasons last (2016) year was 17. Management has made improvements to safety this year and presently there record shows 6 accidents. Does this represent a significant change to the safety record?*
- *Specify the problem in probabilistic terms.*
- *Is this a Poisson model? In what way? What would t and λ be? Given that it is Poisson, what are the parameters and probability calculation are we making?*
- *What is the "risk" in this scenario?*

Poisson Distribution - 7

- ***Analysis/answers*** *– It is a Poisson model because accidents can occur at any time and are random events. We have choices for the time units. We could use months, weeks, days even a year. "months" is a common choice for this. So t=12 months (last year) and the rate is λ=17/12=1.417 per month.*
- ***This year X=6 at 10 months.***

- ***If the old rate is still correct, what is the probability P(X≤6)? Use a cumulative Poisson: The mean in 10 months is λt=1.417(10)=14.17***

$$P(X \leq 6) = 0.0128$$

- *We have used a cumulative Poisson and Minitab for this calculation. Since this chance probability is "small", we have good reason to believe that the rate has dropped from its former value.*

Monte Carlo Simulation - 1

- **MC simulation** has an enormous number of methods and applications.
- Essentially, we have several variables and these are connected/related in some way. Each variable is random in some way – this randomness can be assumed, postulated, or modeled from empirical data. There is some desired output that is of interest and this is related to the several input variables. One iteration of the final model gives us one output. Many iterations of the model gives us a distribution for the final output.

Monte Carlo Simulation - 2

- Live example: Monte Carlo for determining a conservative value for a Cpk/Ppk statistic.
-

Likelihood / Consequent Matrix

Likelihood rating						
E	IV	III	II	I	I	I
D	IV	III	III	II	I	I
C	V	IV	III	II	II	I
B	V	IV	III	III	II	I
A	V	V	IV	III	II	II
	1	2	3	4	5	6
	Consequence rating					

LC Matrix example

Hazard Severity Category		Hazard Probability				
		FREQUENT $>1x10^{-4}$	PROBABLE 1×10^{-4} to 1×10^{-5}	OCCASIONAL 1×10^{-5} to 1×10^{-6}	REMOTE 1×10^{-6} to 1×10^{-7}	IMPROBABLE $<1 \times 10^{-7}$
CATASTROPHIC - Safety Critical event resulting in: death; aircraft loss or damage beyond economical repair; or severe environmental damage.	I	HRI 1	HRI 2	HRI 4	HRI 8	HRI 11
CRITICAL- Safety Critical event resulting in: Severe injury or occupational illness to any personnel that results in a permanent partial disability; aircraft or property damage > $1,000,000; a condition that requires immediate action to prevent the above (including Cat I) or major environmental damage.	II	HRI 3	HRI 5	HRI 6	HRI 10	HRI 15
MARGINAL- Minor injury or occupational illness; aircraft or property damage > $10,000; a inflight failure requiring termination of flight for safety reasons or correctable environmental damage.	III	HRI 7	HRI 9	HRI 12	HRI 14	HRI 17
NEGLIGIBLE - Less than minor injury or damage or minimal environmental damage. Mission can be continued with minimum risk.	IV	HRI 13	HRI 16	HRI 18	HRI 19	HRI 20

Hazard Risk Index	Safety Risk	Decision Authority For Residual Risk
1 - 3	HIGH	Component Acquisition Executive
4 - 7	SERIOUS	Program Executive Officer or equivalent
8 - 10	MEDIUM	JPO Program Manager or equivalent
11	LOW	Acceptable Risk with review by JPO PM or designee. Hazard roll-up must not exceed LOA value
12 - 20	VERY LOW	Acceptable

Tools Summary- Risk Planning

- FMEA; PFMEA; DFMEA
- Checklists
- Redundancy
- DFV considerations
- Standards and Standard Work
- Probabilistic Risk
- SIPOC, Process Mapping
- Likelihood/Consequent Matrix

Completion of tools section

Q&A