CMHT
CENTRE FOR MECHATRONICS AND HYBRID TECHNOLOGIES

ENGINEERING | McMaster University

How to select the proper sampling size

Ben Miethig

August 2019

Validating Parts in the Automotive Industry

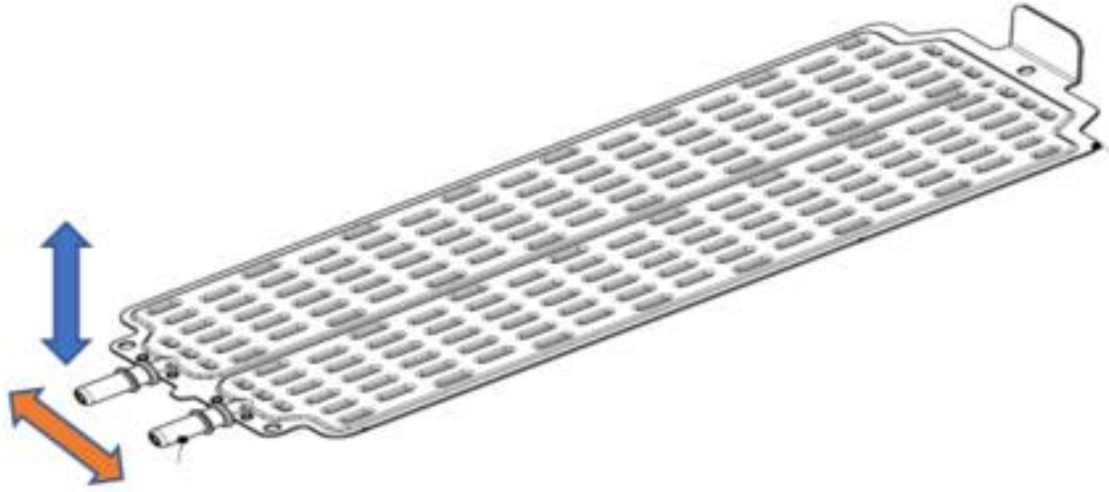
- Quality assurance is critical to ensuring parts manufactured for vehicles meet and maintain design specifications set by the customer
- Quality assurance requires lots of parts to be tested to ensure the number of defective parts falls within specified ranges
- This presentation will outline how to select an adequate sample size for verifying the number of defects in a given lot with statistical significance
- A practical example of tests completed with a glycol battery cooler is used to illustrate the collection and analysis of attribute data to determine the likelihood of cyclical loading causing leakage in the cooler’s fluid fittings

CMHT
CENTRE FOR MECHATRONICS AND HYBRID TECHNOLOGIES

Initial Battery Cooler Test Procedure and Results

ENGINEERING McMaster University

- 10 battery coolers tested by applying a cyclical bending moment to the fittings on the cooler
- All plates are leak tested with a nitrogen-air mixture to ensure no leaks are present before or after the cyclical loading
  - If a part exhibits a leak rate greater than 0.4 CCMs, it will be deemed to have failed
- After applying 2000 cycles in both the blue and orange directions on both fittings of all 10 plates, no failures were found to occur
- What does this mean statistically?



Statistical Relevance – The Binomial Distribution

- The binomial distribution is a discrete data distribution which can be used to characterize attribute data which takes on discrete values such as pass/fail
- Classical example is the odds of rolling a certain number on a six-sided dice in a given number of rolls
- All events for this distribution are independent, for example, rolling a five on a die will no impact on a five being rolled on the same die in a subsequent roll
- This can be extrapolated to manufacturing settings when determining the defective rate of parts in a lot/population

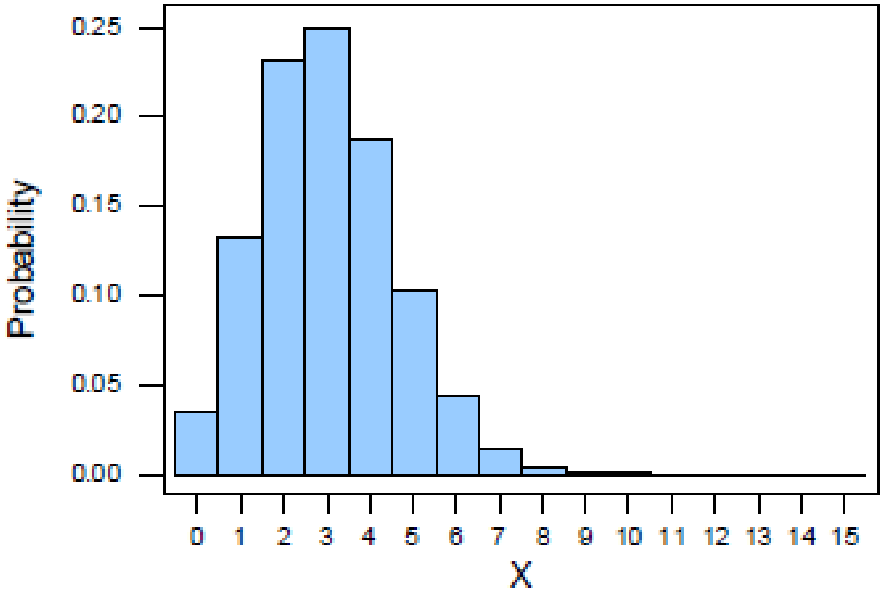
The Binomial Distribution

- The binomial distribution can be used to determine the probability of an event occurring a given number of times in a sample size or sequence of events

$$P(a,n) = \binom{n}{a} * p^a * q^{n-a}$$

$p = probabilty\ of\ an\ event\ occurring;$
$q = probability\ of\ an\ event\ not\ occurring = 1 - p;$
$n = number\ of\ samples;$
$a = number\ of\ events\ that\ occurred;$

Binomial distribution with n = 15 and p = 0.2



[https://newonlinecourses.science.psu.edu/stat414/node/70/]

The Poisson Distribution

- The Poisson distribution is a special case of the binomial distribution where the sample size, n, is large, and the probability of an event occurring, p, (a defective part in this case) is low:

$$P(a,\mu) = \frac{e^{-\mu}\mu^{a}}{a!}$$

$where\ \mu = mean\ number\ of\ events\ in\ a\ given\ quantity = n * p;$
$e = base\ of\ natural\ log;$
$p = probabilty\ of\ an\ event\ occurring;$
$q = probability\ of\ an\ event\ not\ occurring = 1 - p;$
$n = number\ of\ samples;$
$a = number\ of\ events\ that\ occurred;$

Example 1 – Statistical significance of testing 10 parts

- 10 parts is a very small sample size so conclusions about the population of parts cannot be made with much confidence
- Need to know the % confidence interval and sample size, conclusions can be drawn for a new sample of parts taken from the same population
- The % confidence interval is the degree to which the conclusions can be trusted. Higher % means there is more certainty in the drawn conclusions
- Assume a 90% confidence interval with a sample size of 10 and no defective events occurring (a=0)

Example 1 cont.

- From the table, when a=0, n=10, and confidence interval = 90% (reflected in the upper limit of 0.1 (100%-90%))
  - Upper confidence limit = 2.06
- This means that in a new sample of 10 parts from the same population, 2.06 parts could be expected to fail, corresponding to a failure rate of 20.6%

a	n	$p = \frac{a}{n}$	Confidence limits for $\mu = np$ corresponding to probabilities of:		
			Upper limit		
			0.1	0.025	0.005
0	5	0	1.84	2.61	3.27
	10		2.06	3.09	4.11
	20		2.17	3.37	4.65
	∞		2.30	3.69	5.30

- To bring predicted expected failure rate down, the number of samples must be increased.
  - If samples tested increased to 20 with 0 failures, then the expected number of failures in a new set of 20 parts from the same population would be 2.17, or 10.85%

Table values calculated using the binomial and Poisson distribution formulas. See MATLAB Program for further derivation.

Risk Associated with Accepting Parts with Defects

- Producer risk: the risk that a difference in defects between two lots – or populations – is found when the lots have the same defect rate
  - producers could risk throwing away a perfectly good lot of parts since the sampled parts misrepresent the defect rate in the population
- Consumer risk: the risk that there are more defects in a lot despite sampling results
  - a risk arises when two lots are assumed to have the same defect rate when they actually have different amounts of defects and this risk decreases as more parts are sampled
- Since it is very costly to sample many parts, producers and consumers must agree on sampling plans with a certain degree of confidence that all parts in a population will meet the desired level of quality

Example 2 – Selecting Sample Sizes

- **Assume a car manufacturer is willing to accept a 10% chance they will receive a lot of parts with defects in 1% of them. What sample size would be required by the consumer to confirm that this criteria has been met?**
- Assume the sample size of parts is large and one defect is found, then the upper confidence limit is 3.89 (3.89 parts or less in the sample will be found defective)
- from $\mu = n * p$:
  - $n = \frac{\mu}{p} = \frac{3.89}{0.01} = 389$

a	Confidence limits for $\mu = np$ corresponding to probabilities of:					
	Lower limit			Upper limit		
	0.005	0.025	0.1	0.1	0.025	0.005
0	0	0	0	2.30	3.69	5.30
1	0.005	0.025	0.105	3.89	5.57	7.43

- Therefore the manufacturer will have to inspect 389 parts to ensure the consumer’s specifications have been met with 90% confidence

Conclusion

- It is not as simple as testing an arbitrary number of parts from production for failure
- To reliably verify parts meet required quality standards, the binomial or Poisson distributions must be used with sampling plans to ensure suppliers and consumers have transparent expectations on part quality
- For further information, please consult the corresponding case study for this presentation or Statistical Methods in Research and Production by O. L. Davies and P. L. Goldsmith (1977).
- For mathematical derivations of the tables used, please view the accompanying MATLAB script.

Acknowledgements

- A special thank you must be extended to John Burgers and Dana Canada Corporation. Without their help and support, this case study would not have been possible. Funding from the NSERC CREATE program is also acknowledged in making this case study possible.

- O. L. Davies and P. L. Goldsmith, Statistical Methods in Research and Production, 4 ed., Edinburgh, Scotland: Longman Group Ltd., 1977.
- Stat Trek, "Poisson Distribution," Stat Trek, 2019. [Online]. Available: https://stattrek.com/probability-distributions/poisson.aspx. [Accessed 1 August 2019].