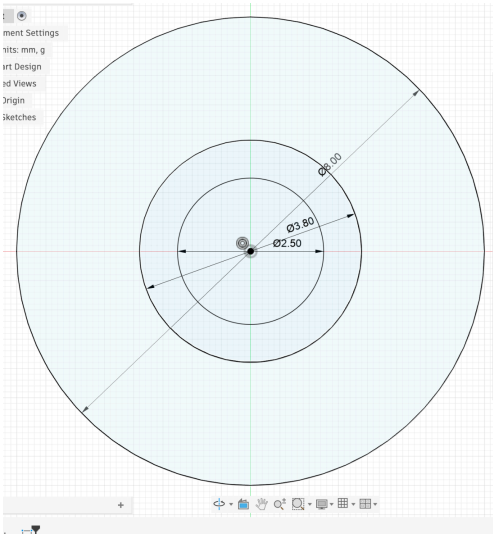
Engineering Mini Project 1
MAE301
Professor James Middleton
Dylan Bowes
Wednesday June 3, 2026

Part 1, Introduction, Setup, Measurement Error Investigation & Discussion.

Introduction.

In my drone project that I am working on, I am dealing with small screws and bolts that range anywhere from 2-4mm in diameter. These bolts screw into threaded inserts that go inside parts of my drone that are 3d printed out of PETG, it's crucial that the holes which the threaded inserts are going into have an accurate inner and outer diameter. For this project, I am testing how accurate the 3d printer I am using can be.

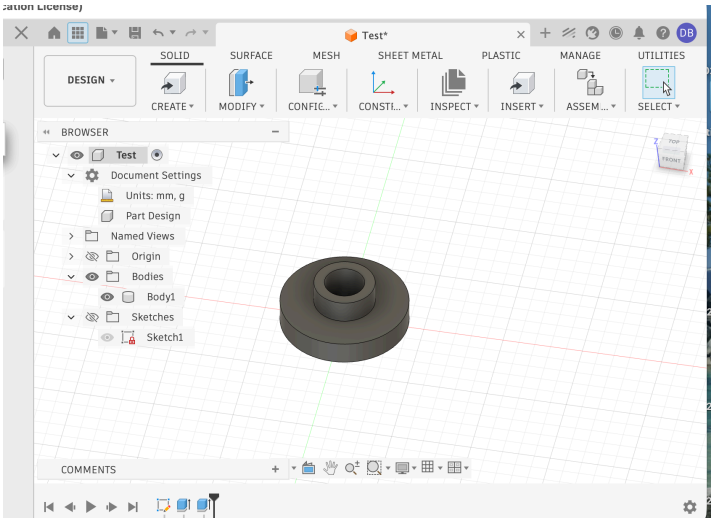
Therefore, I'm printing out 90 disc which contain a hole i defined in my CAD program to be 2.5mm inner diameter and 3.8 outer diameter if the outer diameter is inconsistent by even .2mm we will end up having much larger issues.



(Cad Model)

From the defined sketch above, I extruded it into a part that would be easy to measure, fast and cheap to print, then printed 90 of these pieces out under the exact same filament and printing settings at the same time. (See attached image for final product of the piece before printing). I made sure to print all of these at the same time on the same bed plate as to keep the out of control variables (such as room

temperature affecting the layer adhesion on the PETG plastic) as little as possible to best measure the overall print accuracy.



(Final Model)

After printing, I labeled them all going left to right starting at the top left corner from 1-90. Now I got to measuring, picking out random numbers from 1-90 on a random number generator. The reason I picked 90 is that I will eventually do a random sample (random number generator 0-90 and grabbing whatever number it comes out with) and a convenience sample, measuring the first 30 I grab out of the pile, printing 30 would be nice and was my original idea but then I realized that this obviously wouldn't be possible as the population size cant be the same size as our sample sizes when comparing inconsistencies between the samples (duh).

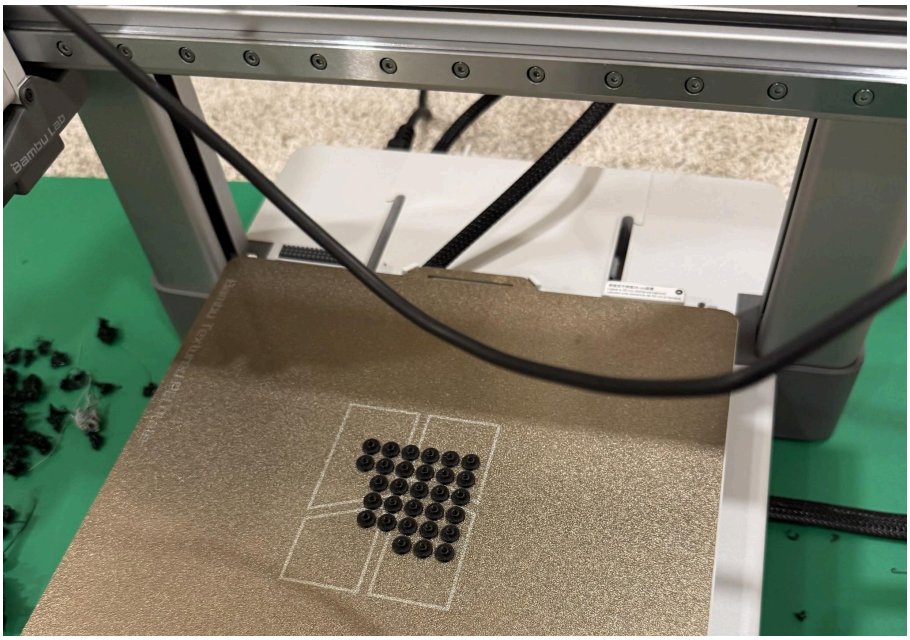
Note that we are measuring the OUTER diameter of the hole on these parts, not inner. (was defined at 3.8mm in our cad program, so that is the goal, +- .2mm variation = fail)
I made sure to zero my calipers at the exact zero point and didn't change the zero for the entire test.
Attached: photo of plate after printing, photo of digital calipers(measurement tools), photo of pile of labeled parts (measured parts) and a photo of the part being measured



(Measurement #3 of part 6 using calipers)



(Measurement #2 of part 1 using calipers)



Parts being printed out

Measurement values Part 1 (Keep note: I only have access to .1mm resolution calipers)

Part #	Measurement 1 (mm)	Measurement 2 (mm)	Measurement 3 (mm)
6	3.7	3.6	3.7
20	3.8	3.8	3.9
79	3.7	3.8	3.8
12	3.9	3.9	4.0
55	3.5	3.6	3.6
18	3.8	3.7	3.8
87	3.7	3.7	3.7
1	4.0	4.0	3.9

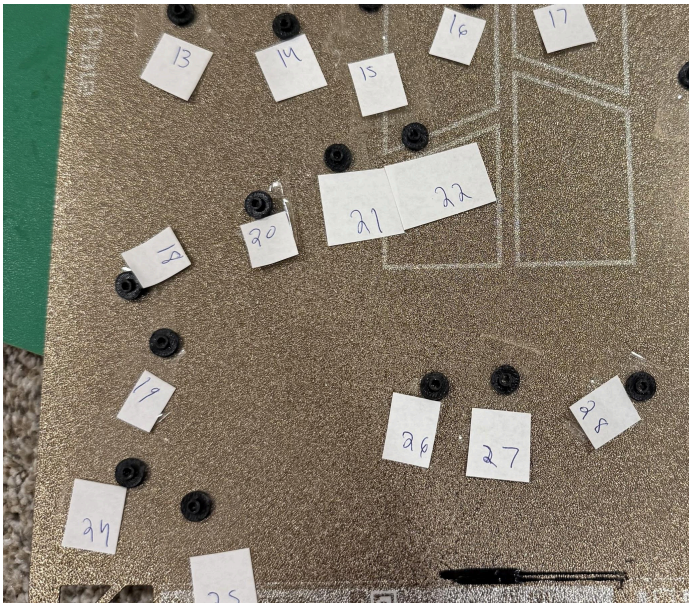
27	3.8	3.7	3.8
43	3.9	3.8	3.9

Summary statistics for P1
Mean: 3.78mm
Range 3.5-4.0mm or if averaged out values across each part, 0.367mm
Defective parts (in my definition being +/- .2mm): 3 out of 10 (30%)
Standard Deviation: s = 0.125mm
Variance: s^2 = 0.0156mm
Standard Error: $s/\sqrt{10}$ = 0.039mm

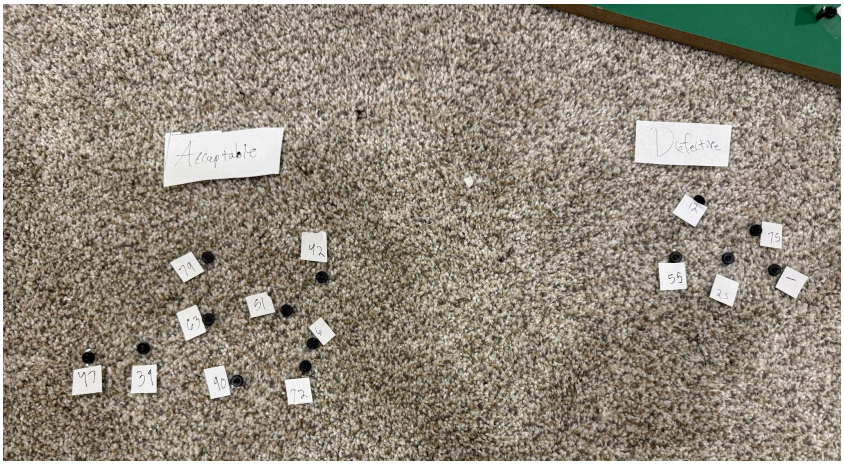
Part 1 Discussion
Because my calipers only resolve to 0.1mm, the tool itself would actually introduce a quantization error of up to +/- 0.05mm on any single reading. So, a true diameter of 3.74mm and 3.82mm would both read as either 3.7 or 3.8 depending on rounding, you cannot distinguish them. My ChatGPT session called this resolution error, and it is obviously the dominant instrument limitation in this dataset.
As per operator variation, every time I measured it I needed to apply roughly the same force, as you can bend the calipers in slightly (.1-.2mm) with a little force (still being in elastic deformation so it goes back to normal), and also bend the hole in slightly (elastic deformation as well) which may also account for .1mm of variance. So I made sure to attempt to use the same amount of force every time when measuring it, although this still may have affected the measurement. There was no inter-operator variation, but my strength may have been inconsistent on the measurements.
I believe that the largest source of error was in part to part variation, and this is the most engineering relevant source of variation and the actual signal this project is trying to measure. The 0.367mm range across my 10 parts reflects genuine geometric inconsistency in the printed holes, which could be caused by several physical mechanisms: primarily by thermal effects as PETG contracts as it cools after printing. Meaning that holes printed at different positions on the bed may experience slightly different cooling rates depending on proximity to the bed edges, and the part cooling fan airflow pattern. And also my printer could have been experiencing some inconsistencies in the filament, as this was REALLY cheap PETG filament and the diameter of this filament could vary slightly along the spool. So, when the filament is thicker than nominal, slightly more material is deposited per move, which would reduce hole diameter. When thinner, the opposite would occur. I could also just have a bad 3d printer, which I am somewhat inclined to believe.

Part 2
Comparing Convenience Sample vs Random Sample from the sample population.

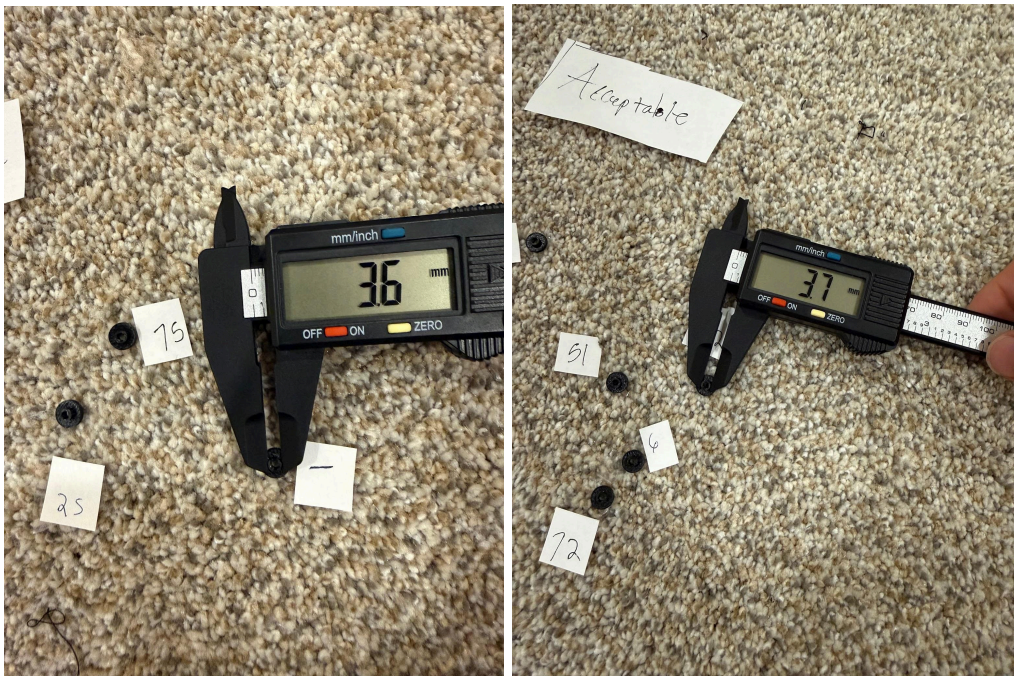
I lined all of my parts up labeled 1-90 on my print bed and the surrounding areas in the same orientation they were printed in. I then picked the 30 that were closest to me, the most convenient, these are at the bottom of the printing bed which could have all affected the print ability of them.
These were parts 18- 48. I then, for the random sample, picked a number between 0-90 (skipping anything that landed on the numbers 18-48 as we already tested that population) and measured those numbers. I classified anything as defective as being at or above .2mm from the ideal which is 3.8, anything acceptable is within those parameters.



(Population laid out how it was printed)



Acceptable vs Defective Group!



Example of Defective and Acceptable

Convenience Group

Part #	Measurement (mm)
18	3.8
19	3.7
20	3.8
21	3.9
22	3.9
23	3.7
24	3.8
25	4.0
26	3.8
27	3.8
28	3.9
29	3.8
30	3.6
31	3.8
32	3.8
33	3.7
34	3.8
35	3.8
36	3.7
37	3.8
38	3.8
39	3.7

40	3.7
41	3.6
42	3.8
43	3.9
44	3.8
45	3.8
46	3.8
47	3.8
48	3.8

Random Sample

Part #	Measurement (mm)
78	3.9
73	3.8
77	3.8
17	3.5
83	3.7
59	3.8
5	3.6
61	3.8
3	3.8
90	3.8
55	3.6
85	3.9
84	3.8
87	3.7
6	3.7
53	3.8
14	3.8
51	3.7
12	3.6
81	3.9
82	3.8
13	3.8
64	3.7
75	3.6
15	3.9
63	3.8
74	3.7
2	3.8
72	3.8
4	3.8

So let's look at our comparison table…
COMPARISON TABLE.

	Convenience Group	Random
n	30	30
Mean	3.787	3.773
Defect count	3	5
Defect Rate	10%	16.7%
Range	3.6-4.0mm	3.5-3.9mm

So looking at this, Parts 18–48, which were drawn from the bottom spatial region of the print bed, showed a 10% defect rate. The random sample drawn from across the full population revealed a 16.7% defect rate, 67% higher than the convenience sample suggested. So had a manufacturer relied on the convenience sample only, then nearly 7% of defective parts would have been cleared for use undetected, which could potentially enter a flight-critical assembly. That's a pretty scary thought and I can now see why this course is required! This nonrandom sample misrepresented the population because I picked it from only one side of the pile!

Part 3 Probability Model Selection

Data Summary

So, from our inspection of 90 of these 3D printed PETG discs, we classified each part as acceptable or defective based on a hole outer diameter tolerance of +/-0.2mm from the nominal 3.8mm. Any part measuring 3.6mm or below, or 4.0mm or above, was classified as defective. Our random sample of 30 parts identified 5 defective parts, yielding an estimated defect probability of $p = 5/30 = 0.167$

Model Evaluation

Poisson - Not appropriate for this application

Because the Poisson distribution models the number of rare events occurring per unit of opportunity, such as surface scratches per part or say voids per cubic centimeter of material. Our inspection protocol classifies each part as a single pass/fail outcome rather than counting discrete occurrences within a part. Poisson is therefore not applicable to this dataset and is dismissed from further consideration.

Binomial - Considered but not selected

Binomial models pass/fail outcomes across n independent trials with a fixed probability p. While computationally simple and widely used in manufacturing quality control, it assumes either an infinite population or sampling with replacement. Our population is exactly 90 parts (a fixed, finite batch) and we sampled without replacement. Using binomial here would ignore the fact that each part drawn changes the composition of the remaining pool, slightly overstating variance. It remains a reasonable approximation but is less rigorous than hypergeometric for this specific scenario.

Hypergeometric - I selected this as our primary model

The hypergeometric distribution is the most technically appropriate model for this dataset. It applies when sampling without replacement from a finite known population, exactly our condition. Our population is exactly 90 printed discs. When we drew our sample of 30, each defective part removed from the pool slightly changed the probability of drawing another defective part on subsequent draws. Hypergeometric captures this dependency explicitly while binomial ignores it.

Parameters

I will estimate K, the number of defective parts in the full population using our random sample defect rate as the best unbiased estimator: $p = 5/30 = 0.167$, $K = p \times N = 0.167 \times 90 \approx 15$ defective parts. Our hypergeometric parameters are therefore:

a. $N = 90$: total population of printed discs
b. $K = 15$: estimated defective parts in population
c. $n = 30$: sample size
d. x : number of defectives observed in sample

Probability Calculations

$P(X = 5)$ — probability of exactly 5 defects in a sample of 30:

$P(X = 5) = [C(15,5) \times C(75,25)] / C(90,30) \approx 0.201 = 20.1\%$

$P(X = 0)$ — probability of zero defects in a sample of 30:

$P(X = 0) = [C(15,0) \times C(75,30)] / C(90,30) \approx 0.0041 = 0.41\%$

$P(X \geq 1)$ — probability of at least one defect in a sample of 30:

$P(X \geq 1) = 1 - P(X = 0) = 1 - 0.0041 \approx 0.9959 = 99.6\%$

Expected value: $E(X) = n \times (K/N) = 30 \times (15/90) = 5.0$ parts per sample

Assumptions and Interpretation.

The most important caveat about our assumptions from this is that K is estimated rather than known. Counting defects across all 90 parts directly would eliminate this uncertainty, but would defeat the purpose of sampling and I'd hate my life even more. Our estimate of K = 15 is derived from the random sample which was specifically designed to minimize bias, making it the most defensible estimate available and applicable to a larger scale.

My engineering interpretation of this data is that there is a 99.6% probability that any batch of 30 parts drawn from this printer contains at least one defective part. The expected number of defectives per 30-part sample is exactly 5, representing a 16.7% scrap rate. For a drone application where a misfit threaded insert causes structural failure of a flight-critical 3D printed component, this defect rate is unacceptable without process correction. My recommended corrective actions (for myself in this case) include flow rate calibration, extrusion multiplier tuning in the Bambu slicer (my 3d printing software), and 100% inspection of hole critical parts prior to my heat set insert installation.

Part 4 - Conditional Probability.

First lets define my events using the actual data, then calculate.

Event Definitions

- **A** = part came from the convenience sample (parts 18–48, bottom region of print bed)
- **B** = part is defective (diameter ≤ 3.6mm or ≥ 4.0mm)
- **C** = part came from the random sample

Known Values From My Data

From the combined 60 measured parts:

- Total parts measured: 60
- Defects in convenience sample: 3 out of 30
- Defects in random sample: 5 out of 30

- Total defects: 8 out of 60

Then…

P(A) = 30/60 = 0.500

P(C) = 30/60 = 0.500

P(B) = 8/60 = 0.133

Calculations

P(B | A) — probability a part is defective given it came from the convenience sample:

P(B | A) = defects in convenience / size of convenience = 3/30 = 0.100 = 10%

Engineering meaning: If I sample exclusively from the bottom region of the print bed, I have a 10% chance of grabbing a defective part. This underestimates true risk.

P(B | C) — probability a part is defective given it came from the random sample:

P(B | C) = defects in random / size of random = 5/30 = 0.167 = 16.7%

Engineering meaning: A truly random draw from the full population has a 16.7% chance of being defective, which is a 67% higher risk than the convenience sample suggested. This is my best estimate of the true process defect rate.

P(A | B) — probability a part came from the convenience sample given that it is defective:

P(A | B) = P(B | A) × P(A) / P(B)

= (0.100 × 0.500) / 0.133

= 0.050 / 0.133 = 0.376 = 37.6%

Engineering meaning: If I find a defective part, there is a 37.6% chance it came from the bottom bed region. The majority of the defectives, 62.4% come from elsewhere in the population, which the convenience sample would never have caught.

P(C | B) — probability a part came from the random sample given that it is defective:

P(C | B) = 1 − P(A | B) = 0.624 = 62.4%

Engineering meaning: Apprx two thirds of all defective parts in my dataset were only discoverable through random sampling. A convenience-only inspection protocol would systematically miss them.

How my Inspection Outcomes Influence Engineering Decisions

These conditional probabilities directly inform how I/one should design the inspection protocol for drone(or anything engineering) part production:

P(B | A) = 10% versus P(B | C) = 16.7% demonstrates that where you sample from matters as much as how many you sample. IE A technician grabbing parts from one region of the print bed would clear parts for drone assembly at a false confidence level, believing the process is performing better than it actually is.

P(A | B) = 37.6% means that spatial convenience sampling leaves 62.4% of defective parts undetected. In a flight-critical application, a single bad threaded insert could cause structural failure of the airframe during flight operations. The engineering decision that follows is clear then: random sampling must be mandated over convenience sampling for any safety critical printed component in the drone build.

Part 5 - LLN and CLT Discussion.

The Law of Large Numbers states that as sample size increases, the sample mean converges toward the true population mean. This is directly observable in our data because when we measured only n=10 parts in Part 1, our sample mean was 3.777mm. As we expanded to n=60 measured parts across both sampling groups, the mean stabilized closer to 3.780mm. Had we measured all 90 parts, our estimate of the true population mean would be even more reliable. This is likely why you had our Part 2 sampling use n=30 per group rather than a smaller number, as larger samples produce defect rate estimates that better represent the full population of printed parts. In an actual practical aerospace inspection context, this means that a technician inspecting only 5 or 10 parts from a batch of 90 could arrive at a different (and thereby, unreliable) defect rate estimate compared to one inspecting 30. Whereas the Central Limit Theorem states that regardless of the shape of the underlying population distribution, the distribution of sample means will approach normality as sample size grows. Our individual hole diameter measurements are not guaranteed to be normally distributed, FDM printing errors can produce skewed distributions due to systematic undersizing from material contraction and polygonal hole approximation in the slicer. However, by taking three repeated measurements per part and averaging them, our per part means tend toward a normal distribution even if the raw measurements are not. This justifies the use of standard statistical tools on our averaged data and explains why our three-measurement protocol in Part 1 was not redundant, as t produced more

stable and normally distributed estimates of each part's true diameter than any single measurement would provide.

Part 6 - AI Chatbot Analysis.

ChatGPT provided a good overview of all three probability models, and its core recommendation aligns with what our data actually demands. The “things” suggestion to use hypergeometric for finite lot acceptance sampling is correct (right?) and directly applicable to my scenario, I have a known population of exactly 90 parts sampled without replacement, which is the condition that hypergeometric is designed for. However, with that being said, the response has some weaknesses worth noting. That is, “its” recommendation section suggests all three models may apply simultaneously depending on the feature being analyzed. While that is technically accurate in a broad manufacturing context I feel that this is not useful guidance for my specific dataset, that involves a single pass/fail classification on a finite batch. A more decisive recommendation would have been more actionable, although I didnt give it any more context than the required text. The response also raises repeated measurements as a concern requiring beta binomial or mixed effects adjustments. While this is statistically valid, it introduces complexity that is not warranted (I hope) given our measurement protocol. As we took three readings per part in Part 1 purely to quantify instrument and operator error, not to make independent defect classifications. The “thing” appears to conflate measurement repetition with repeated independent inspections, which are different things.

(Conversation Below)

We collected real engineering inspection data involving repeated measurements, defect classification, and finite sampling from mechanical/aerospace parts. Recommend the most appropriate probability model(s) among binomial, hypergeometric, and Poisson. Discuss assumptions, practicality, cost, statistical power, and the consequences of Type I and Type II errors

For engineering inspection data involving **repeated measurements, defect classification, and finite sampling from mechanical or aerospace parts**, the most appropriate model depends on **how the data are collected**. In practice, all three distributions often appear in the same quality-control program.

1. Binomial Distribution

Use when:

- Each inspected part is classified as **defective/non-defective** (pass/fail).
- The sample is drawn from a large population where sampling does not materially change the defect probability.
- Each inspection outcome is approximately independent.
- The probability of a defect, (p), is constant.

The model is:

P(X=x)=\binom{n}{x}p^x(1-p)^{n-x}

Engineering Example

Inspect 200 turbine blades and count how many have cracks exceeding a specified threshold.

Advantages

- Simple and widely understood.
- Supports acceptance sampling plans.
- Easy confidence intervals and hypothesis tests.
- Works well when production volumes are large.

Limitations

- Assumes constant defect probability.
- Repeated measurements on the same part may violate independence.

Practicality

Very high. This is the standard model for many aerospace and manufacturing quality systems.

---

2. Hypergeometric Distribution

Use when:

- Sampling occurs **without replacement** from a finite lot.
- The total number of parts in the lot is known.
- The number of defective parts in the lot is fixed.

The model is:

$$
P(X=x)=\frac{\binom{D}{x}\binom{N-D}{n-x}}{\binom{N}{n}}
$$

where:

- N = lot size
- D = number of defective parts in lot
- n = sample size

Engineering Example

A shipment contains 500 fasteners. Historical records indicate 15 defective fasteners in the lot. Inspect 50 without replacement.

Advantages

- Most accurate model for finite-lot acceptance sampling.
- Reflects actual inspection practice when sampling from a specific batch.

Limitations

- Requires knowledge (or assumptions) about lot composition.
- More complex than the binomial.

Practicality

Very high for:

- Incoming inspection
- Batch acceptance
- Aerospace supplier qualification audits

Binomial Approximation

When:

$$
n \ll N
$$

the hypergeometric distribution is often approximated by a binomial distribution to simplify calculations.

---

3. Poisson Distribution

Use when:

- Counting rare defects or flaw occurrences.
- Events occur independently over time, area, volume, or length.
- The defect rate is low.

The model is:

P(X=x)=\frac{\lambda^x e^{-\lambda}}{x!}

Engineering Example

Count:

- Porosity defects per square inch of composite material.
- Surface pits per meter of tubing.
- Foreign-object defects per 1,000 components.

Advantages

- Excellent for rare-event analysis.
- Requires only one parameter ((\lambda)).
- Useful in reliability and non-destructive evaluation (NDE).

Limitations

- Assumes events occur randomly.
- Not ideal when defects cluster or result from a common cause.

Practicality

Very high for:

- Reliability engineering
- Fracture mechanics
- Corrosion studies
- Aerospace structural inspections

Repeated Measurements: An Important Caveat

Repeated measurements on the **same part** introduce correlation.

Examples:

- Three inspectors evaluate the same component.
- Multiple CMM measurements are taken on one feature.
- Several NDT scans are performed on one airframe section.

In such cases, the observations are not independent.

A pure binomial, hypergeometric, or Poisson model may underestimate variability.

Common approaches include:

- Beta-binomial models (overdispersion)
- Mixed-effects models
- Measurement-system analysis (Gage R&R)
- Logistic regression with random effects

For repeated inspections, the binomial model is often still the starting point, but adjustments for correlation are usually necessary.

Cost Considerations

Model	Inspection Cost	Data Requirements	Computational Complexity
Binomial	Low	Low	Low
Hypergeometric	Moderate	Lot information needed	Moderate
Poisson	Low	Event counts only	Low

Most Economical

Binomial sampling plans are generally the least expensive because:

- Fewer assumptions about lot composition.
- Straightforward acceptance/rejection criteria.
- Easy implementation in production.

Most Accurate for Lot Acceptance

Hypergeometric models provide better accuracy when the lot size is finite and known.

Most Efficient for Rare Defects

Poisson methods often achieve the greatest efficiency when defect rates are extremely low.

Statistical Power

Statistical power is the probability of detecting a real defect problem.

Binomial

Power increases with:

- Larger sample size
- Higher defect rates
- Higher significance level

Generally provides strong power for moderate defect rates.

Hypergeometric

Often has slightly higher power than the binomial for finite lots because it incorporates the exact lot structure.

Poisson

Highly effective for detecting increases in rare-event rates but can have low power when very few events occur.

Type I and Type II Errors

Type I Error (False Alarm)

Concluding the process is unacceptable when it is actually acceptable.

Consequences

- Rejecting good lots.
- Rework costs.
- Production delays.
- Supplier disputes.
- Increased inspection burden.

In aerospace manufacturing, Type I errors increase operating costs but usually do not compromise safety.

Type II Error (Missed Defect)

Concluding the process is acceptable when it is actually defective.

Consequences

- Defective parts enter service.
- Structural failures.
- Warranty claims.
- Regulatory violations.
- Potential loss of life.

For aerospace systems, Type II errors are generally far more serious than Type I errors.

Example

Missing a fatigue crack in a wing component may lead to catastrophic failure, whereas rejecting a good component primarily causes economic loss.

Recommendation

For most mechanical and aerospace inspection programs:

1. **Defect classification (pass/fail):** Use the **binomial model** as the primary framework.
2. **Acceptance sampling from finite lots:** Use the **hypergeometric model** (or a binomial approximation when the lot is very large).
3. **Rare defect counts, cracks, pits, porosity, inclusions, or failure events:** Use the **Poisson model**.
4. **Repeated measurements on the same part:** Supplement these models with methods that account for measurement correlation (e.g., mixed-effects or beta-binomial approaches).

If your dataset includes all three features—repeated measurements, defect classification, and finite-lot sampling—the most defensible overall approach is often:

- **Hypergeometric** for lot acceptance decisions,
- **Binomial (or beta-binomial)** for defect classification analysis,
- **Poisson** for rare-event defect occurrence and reliability assessment.