Proceedings of the Annual General Donald R. Keith Memorial Conference
West Point, New York, USA
May 4, 2023
A Regional Conference of the Society for Industrial and Systems Engineering

Analysis of Statistical Sampling Procedures at an Army Ammunition Plant

Ryan Santoro, Nicole Sacchinelli, Michael Rocha, and Andrew Cammack

Department of Mathematical Sciences
United States Military Academy
West Point, NY 10996

Corresponding author's Email: ryan.santoro@westpoint.edu

Author Note: We would like to thank Doug Dillon, Patti McGuire, Laura Baumann, LTC Christopher Denton, Adam Hinsdale, and everyone else at the Lake City Army Ammunition Plant for all of their help and support. The views expressed herein are those of the authors and do not reflect the position of the United States Military Academy, the Department of the Army, or the Department of Defense.

Abstract: The United States Army plans to adopt more expensive ammunition in the forms of .50 caliber and 6.8 mm ammunition. This analysis aims to determine if the Army can meet its quality control objectives while reducing the number of rounds destroyed during acceptance testing using .50 caliber ammunition. Incorporating the more accurate hypergeometric equation, quality assurance data, and simulations, this analysis proposes more structured destructive quality assurance testing processes while quantifying associated risks for each model. The two methods to achieve these goals were to change either the sample size or the lot size. Supplying examples with specific sample or lot sizes, this analysis provides the range of risk which could be accepted following the recommended procedure.

Keywords: Small Caliber Ammunition, Acceptance Sampling, Hypergeometric Distribution, Systematic Sampling, Simulation

1. Introduction

As an army ammunition manufacturing plant, Lake City Army Ammunition Plant (LCAAP) manufactures millions of rounds a day. However, with the numerous amounts of rounds flowing through automated machines, quality assurance can prove to be a challenge especially when the United States military is the main client. As a result, military standards were created to provide the baseline quality of products the Department of Defense (DoD) would accept. These standards provide a set of sampling plans and procedures for planning and conducting inspections to assess quality and conformance to contract requirements.

Within the past few years, LCAAP has seen some upgrades in its .50 caliber machinery to include new cartridge loaders which utilize automatic image recognition software. These upgrades have relegated the "legacy" machines to a rarely used status due to the higher efficiency and lower failure rate of the newer machines. With these more recent upgrades in technology, it is important to reevaluate the standards which were originally based on the older equipment. Exploring different distributions or methods of sample selection would provide a look into what others are doing to ensure their quality criteria are being met and compare it to the current DoD system. If the comparison shows that current DoD standards are outdated or are restricting current processes unnecessarily, decision-makers should consider amending the standards. This paper will explore whether the Army can meet its quality control objectives at Lake City Army Ammunition Plant while reducing the number of rounds destroyed during acceptance testing.

2. Background

As the standard cartridge for .50 caliber, this paper will analyze the quality control objectives for the M33 .50 Caliber Ball Cartridge. M33 rounds are purchased by the Department of Defense in batches known as lots. Within each lot which is around 360,000 rounds, there are twenty sublots which are further divided into six trucks. There are four machines that print the rounds. The current procedure at LCAAP to sample rounds for acceptance testing involves a worker selecting a number of rounds from the top of a truck and setting them aside as the samples for destructive testing. Simultaneously, nondestructive sampling takes place in the same manner. However, nondestructive testing is performed immediately for each sublot rather than being held by lot. The rounds set aside for destructive testing are then taken to another building where they undergo several tests to measure their effectiveness.

2.1. Sampling Techniques

One of the hardest things to achieve in studies is having a fully random sample. Bias is an underlying factor in almost every aspect of human choice. One way to get a completely random sample is by having a computer choose and randomly pick numbers. To have a more reliable study and support valid inferences it is required to have a completely random sample (Dobson, Woller-Skar, Green, 2017). Having random samples means that all of the units within the population have the same probability of being picked.

When faced with the prospect of sampling, quality control samplers are faced with two broad strategies they can employ: probability sampling and non-probability sampling. Within each strategy are sub-strategies which each have their separate pros and cons. This paper focuses on probability sampling because it attempts to replicate true random sampling which provides a higher standard for acceptance.

2.1.1. Other Army Sampling

One study performed on Army procedures is the Review of Department of Defense Test Protocols for Combat Helmets. The purpose of the study was to see if combat helmet testing procedures, and ultimately the binomial calculations, are still useful. The committee examined the differences between the true operating characteristic (OC) curves and OC curves found via the assumption that the penetration probabilities are equal across all shots (National Research Council (U.S.), 2014). The study found that the differences were negligible for the range of penetration probabilities and deviations that were relevant to the testing (National Research Council (U.S.), 2014). After testing their query, the study found that the OC curves computed under the assumption of constant probability provide very good approximations.

2.1.2. Systematic Sampling

Many factories use systematic sampling in order to get a representation of the population. According to the article *Procedures and Machinery Used in Pharmaceutical Sampling*, "this method of sampling is a systematic approach that takes a sample out of particular intervals. The intervals determined in this method can be either a time in the production line or a number of products segregated from the output," (Accupack Engineering, 2022). This means that if we want to sample 100 rounds out of a 1000-round truck then we would sample every tenth round that the machine produces. This provides a more accurate representation of the total population in the truck (Qualtrics, 2023).

This way of sampling is resistant to bias because once the starting point is picked, then the samples chosen after that point is truly random since they would iterate over the whole population. However, one limitation of this method is that the population size needs to be known (Qualtrics, 2023).

2.2. Distributions Behind Probability Sampling

2.2.1. Hypergeometric Distribution

While the binomial distribution, which was used in older military standards (MIL-STD-105E, 1989), is only the approximate probability model for successes from a finite dichotomous population due to sample replacement, the hypergeometric distribution is the exact probability for the number of successes in the sample (Devore, 2012). This is due to the difference in replacement. The binomial distribution is an approximation because once the item is drawn, the binomial distribution places the item back into the pool to possibly be sampled again while the hypergeometric distribution sets the item drawn from the pool aside and continues to sample without replacement (Wroughton & Cole, 2013). Therefore, we used the hypergeometric distribution to predict the probability of acceptance using Equation 1. Using x as the number of defects found, n as a set sample size, Np for total number of defectives in the lot, N as the total number of cartridges in a lot, and c as the threshold number of defective items.

$$P(A) = \sum_{x=0}^{c} h(x, n, Np, N) = \sum_{x=0}^{c} \frac{\binom{Np}{x}\binom{N-Np}{n-x}}{\binom{N}{n}} \tag{1}$$

Through the adjustment of c, different graph curves can be created which would provide guidance as to whether a lot should be accepted or rejected as seen in Figure 1. With a set lot size at 360,000 rounds and the number of defects found to be 1 or more, Figure 1 shows the probability of accepting the lot for different sample sizes over a range of true fault rates between 0.01% and 0.3%.

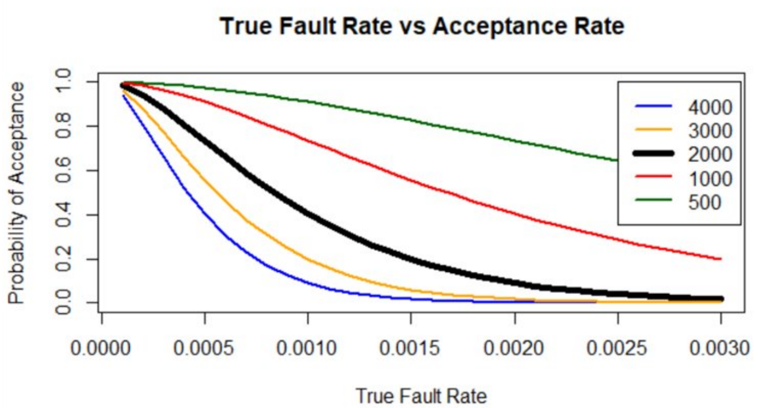

Figure 1. Operating Characteristic Curves

2.2.2. Gamma Distribution

The gamma distribution is used to model data that is right-skewed. According to Dr. Kristen Kuter, "A typical application of gamma distributions is to model the time it takes for a given number of events to occur" (Kuter, 2019). A gamma distribution can be written as followed:

$$f(x;\alpha,\beta) = \frac{1}{\Gamma(\alpha)^{\beta}}(\frac{x}{\beta})^{\alpha-1}e^{-\frac{x}{\beta}} \tag{2}$$

For the gamma distribution function, x is the continuous random variable that represents the rounds that are in the truck. The variable α is the shape parameter. A shape parameter indicates the number of events (Thom, 1958). The variable β is the scale parameter. The scale parameter represents the mean time between events.

2.2.3. Uniform Distribution

A uniform distribution is a continuous probability distribution function where all events under an interval are equally likely to occur. Uniform distribution can be written as followed:

$$p(x) = \frac{1}{b-a} \tag{3}$$

For the uniform distribution function, x is the continuous random variable. Variable a is the lower limit for the interval that is being evaluated. The variable b is the upper limit of the interval that is being evaluated, which will be used in unison with the gamma distribution.

3. Data

The data provided by LCAAP included the defect logs for fiscal years 2014 through November of the fiscal year 2023. The data included the date the lot was tested, the round type, the lot number, the weapon fired, the temperature the ammunition was before firing it, whether the defect was charged to the weapon or ammunition, and whether there needed to be a double sample retest. The data was used to look for patterns, evaluate the tests, and gather the total number of rounds approved for each test given the round type. While cleaning the data, only entries for M33 ammunition in the fiscal year 2022 were kept. One lot

ISBN: 97819384964-4-8

data point was dropped because it claimed a lot size of 3,448,817 rounds. We omitted it under the assumption it was an error or a major outlier. Our analysis found that in the fiscal year 2022, there were 44 lots of 14,142,172 rounds of M33 ammunition that went through the acceptance process with a maximum lot size of 367,548 rounds, a minimum of 107,942 rounds, and an average of 321,413 rounds. Through an assumed true fault rate of 0.005% and the current procedure's sample size of 1945, we calculated the percentages accepted in Table 1 by plugging the given numbers into the hypergeometric equation. From there we set the different ranges and possible lot sizes and sample sizes to find different rates of acceptance.

Table 1. Summary of Data

	Lot Size	Percentage Accepted	True Fault Rate	Extra Risk Accepted	Total Annual Cost
Mean	321,413	99.59521%	0.005%	N/A	$208,216.14
Min	107,942	99.68701%	0.005%	-0.09180%	$619,993.83 (+$411,777.69)
Max	367,548	99.58486%	0.005%	+0.01035%	$182,080.64 (-$26,135.50)

4. Simulations

The above table results were created with the major assumption that the sample was randomly sampled. Rather than accepting this incorrect assumption, we wanted to simulate what the sampling procedure would look like following the current procedure rather than random systematic sampling. As previously mentioned, the way sampling is performed right now is by having a worker go to a truck and pick rounds from the top of the truck to be destroyed for acceptance testing. The rounds are selected from the top of the trucks because once the approximately 3,100 rounds are inside the 2-foot wide, 4-foot long, and 18-inch tall truck, it is impossible to access rounds on the bottom without adding several man hours per truck.

There are a few problems with this method. One is that even if the people picking the ammunition do not think that they have a bias when picking rounds, they do. The second is that since they are picking from a truck that is already loaded, then they are only getting a sample of the rounds that are on the top of the pile. This can cause problems because the rounds chosen are not a fully representative sample of the entire truckload.

4.1. Methodology

As rounds are manufactured, they go onto an assembly line and fall into the truck at the end of its production. Therefore, the rounds that are manufactured first from the machine are at the bottom of the truck. We want to compare the probability of denying a lot with the current method of sampling and a random sampling method. In order for a lot to not be accepted the sampling process needs to find 2 defective rounds.

With this knowledge, we created a simulation that represents the status quo at LCAAP. Using the volume and the average number of rounds in a truck. We calculated the average .50 caliber round per cubic inch in the truck. We then assumed that the furthest down a worker could sample was only 2 inches into the truck. This assumption was based on our team's personal experience with attempting to sample rounds with difficulty reaching rounds below the 2-inch mark. This makes the population size only 344 rounds vs the actual 3,100 rounds within the truck.

We captured what a truck of rounds would look like if defective rounds were printed at the beginning of the population by creating an array of numbers that would represent each round printed by the machine which contained numbers 0-3,099 over a lot-sized printing period as shown in Figure 2.

ISBN: 97819384964-4-8

Proceedings of the Annual General Donald R. Keith Memorial Conference
West Point, New York, USA
May 04, 2023
A Regional Conference of the Society for Industrial and Systems Engineering

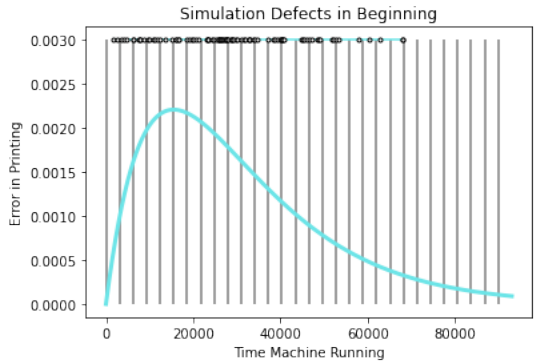

Figure 2. Graph of the Simulation where the Defected Rounds are Printed at the beginning of the Machine's Run Time.

We then used a gamma function to assign a y-value to each round within the truck using the scale parameter 1 with a shape parameter of 4. This created a graph where the peak is at the beginning of the machine's output. Once we achieve the correct graph, we decided to use a uniform distribution from 0 to 1 in order to determine if a round was defective or not. If the y-value from the uniform distribution was smaller than the y-value from the gamma distribution, we labeled it defective which is represented by a 1. If the round was not considered defective then we labeled it with a 0. This process increases the likelihood of a defective round being located near the bottom of the truck or the top of the gamma distribution; while allowing there to be a chance for a detective round later in the printing process.

In order to simulate the current way LCAAP samples .50 caliber rounds, we created four different gamma distributions based on the four machines that print the rounds for a lot. We then took those gamma distributions and split them up into 30 different sections running consecutively to one another to simulate the filling process of all trucks for that machine. Those are the vertical lines seen in Figure 2. We then randomly selected 6 trucks 20 times to simulate the section of the 20 sublots within a lot. However, this number varies based on the size of the lot. Once the trucks were split into the sublots we then sampled 48 rounds from the trucks and determined if they were defective or not based on their value (1 or 0). To simulate the current process we only sampled from the top 344 rounds based on the previous assumption that the person sampling can only reach 2 inches into the truck.

4.2. Results

For the sake of this simulation, in every iteration, the machines printed out more than 2 defective rounds. After running through 10 iterations where the defects were printed at the beginning printing process. We found the smaller the lot size the less likely the lot was successfully denied. This makes sense because there would be fewer rounds to sample from. This causes there to be fewer defects within the machine's printing and it would be much harder to find the defective round since they are only sampling from the top of the truck. These results can be seen in Table 2.

Table 2. Results of the Simulation.

	Lot Size of 111,600	**Lot Size of 334,800**	**Lot Size of 372,000**
Successful	57%	90%	96%
Failed	43%	10%	4%

Based on the results from above with the total annual cost and the accepted risk. We recommend having lot sizes of at least 334,800 to 372,000. This will allow for the annual cost to be less because in proportion they are sampling less than if the lot was smaller. While also upholding a high percent accepted risk since those lot sizes scored well in the simulation. With only failing 4% to 10% of the time. However, we believe if they were to have a truly random sample of the sublots these failure percentages would be even lower than they are now.

5. Conclusions and Future Work

Results from this research indicate that the current process for testing ammunition does not provide the most accurate risk assessment due to a biased sample. Further research should pursue a more accurate data set which may include a full lot being tested to get a better true fault rate. Changes that should be implemented in the current system include set lot size ranges, a better sampling system, and evaluating decision-makers' risk tolerance constraints to see if current procedures meet them. This is further proved through the simulation. It is important to set the sample size within certain parameters so the risk assumed can be calculated with broad assumptions. However, these assumptions cannot be assumed without a fully random sample. A recommendation to improve sampling is to implement truly random sampling, which could include rounds randomly pulled off the line before they get to the truck. The result would be a sample with a more holistic representation of the lot.

Our research recommends lot sizes between 334,800 to 372,000 due to their ability to successfully reject lots that should be rejected. The cost breakdown would also provide a range of costs that decision-makers may find more palatable. Table 3 shows what our solution may look like for risk and cost for a year.

	Lot Size	**Successful Rejection Rate**	**Total Annual Cost**
Min	334,800	90%	$199,890.60
Max	372,000	96%	$179,901.54

Table 3. Recommended Lot Size Ranges

Future research should pursue the refinement of the assumptions of this research. Further exploration of the loading machine fault rates and their effects on sampling procedures. While also looking at different sampling procedures to see which one would be the most accurate while also being inexpensive.

6. References

Baller, D., & Schwartz, B. (2022). Small Arms Ammunition Acceptance Sampling.

Devore, J. L. (2012). Probability and statistics for engineering and the sciences (Eighth edition). Brooks/Cole, Cengage Learning.

Dobson, C., Woller-Skar, M., & Green, J. (2017). Teacher's Toolkit: An Inquiry-Based Activity to Show the Importance of Sample Size and Random Sampling. Science Scope, 040(08). https://doi.org/10.2505/4/ss17_040_08_76

Kuter, K. (2019, March 11). 4.5: Exponential and Gamma Distributions. Statistics LibreTexts. https://stats.libretexts.org/Courses/Saint_Mary's_College_Notre_Dame/MATH_345__-_Probability_(Kuter)/4%3A_Continuous_Random_Variables/4.5%3A_Exponential_and_Gamma_Distributions

MIL-DTL-10190G M33 Spec. (2021).

MIL-STD-105E Sampling Procedures and Tables for Inspection by Attributes. (1989).

MIL-STD-1916 DoD Preferred Methods for Acceptance of Product. (1996).

National Research Council (U.S.), (2014). Review of Department of Defense Test Protocols for Combat Helmets. National Academies Press.

Qualtrics. (2022, January 11). Procedures and Machinery Used in Pharmaceutical Sampling. Accupack. https://www.accupackengineering.com/post/procedures-and-machinery-used-in-pharmaceutical-sampling/

Systematic Random Sampling: The Complete Guide. (n.d.). Qualtrics. Retrieved April 21, 2023, from https://www.qualtrics.com/uk/experience-management/research/systematic-random-sampling/

Thom, H. C. S. (1958). A NOTE ON THE GAMMA DISTRIBUTION. Monthly Weather Review, 86(4), 117–122. https://doi.org/10.1175/1520-0493(1958)086<0117:ANOTGD>2.0.CO;2

Wroughton, J., & Cole, T. (2013). Distinguishing Between Binomial, Hypergeometric and Negative Binomial Distributions. Journal of Statistics Education, 21(1).