7-2007

Testing Naval Artillery and Other Things

Tricia Buchanan
University of Nebaska-Lincoln

Master of Arts in Teaching (MAT)
Masters Exam

Tricia Buchanan

In partial fulfillment of the requirements for the Master of Arts in Teaching with a Specialization in the Teaching of Middle Level Mathematics in the Department of Mathematics.
Gordon Woodward, Advisor

July 2007

Testing Naval Artillery and Other Things That Blow Up

Tricia Buchanan

July 2007

Testing Naval Artillery and Other Things That Blow Up

In WWII a tremendous amount of artillery shells were made to support the war efforts. There were problems with the artillery shells sent to the battlefield; the main problem was their lack of ability to blow things up. In other words, they were duds! While one may think that dud shells were the proverbial rare case, in my paper I hope to show you that instead it unfortunately seemed more the norm. The reasons behind this are varied but in this paper I will focus on the testing practices of the artillery shells and some of the issues that occurred because of this testing.

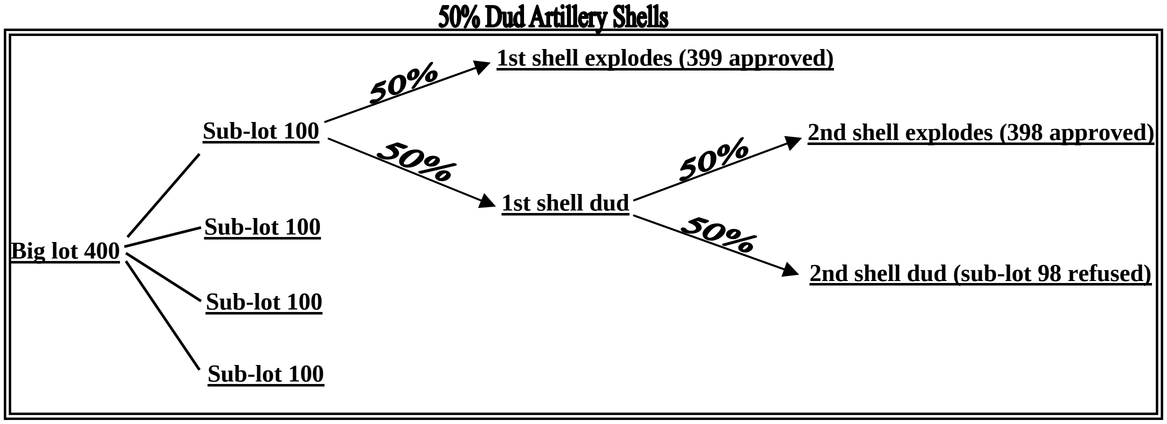
One of the main problems with testing artillery shells is that you are blowing up your supply. If you test each and every shell you definitely will know which shells are duds and which shells explode; however the problem is you have used up your entire stock of shells while doing so. A system had to be devised to test samples of the supplies of artillery shells and then send on what they felt were usable shells into battle. A "system of proof" was developed by the British Ordnance Board.

Artillery shells were first tested in big lots of 400, as described next. The big lots were divided into four sub-lots of 100 shells each. Shells were picked out of the first sub-lot of 100 at random and tested. If the first shell worked (exploded), the entire big lot (now 399) was passed with no further proof needed. If the first shell failed, a second shell was picked and tested. If the second shell worked the rest of the big lot (now 398) was then approved. If the second shell had also failed, the sub-lot of 100 was refused and they started over with the next sub-lot of 100 shells. The shell maker was also given the choice of taking back the entire remaining big lot of 398 shells without further testing, but this rarely happened.

I will be examining several techniques that were developed for the testing of shells and then will discuss the efficiency, or lack there of, later on in the paper. Also, for this paper, we are using the concept of independence for these examples. Two events are independent if the occurrence of one of the events gives us no information about whether or not the other event will occur; that is, the events have no influence on each other. The exploding or non-exploding of the shells has nothing to do with each other. This means that if 50 percent of my shells are duds each shell has a 50/50 chance of being a dud or live shell. Even if I test ten shells in a row that are duds, the following shells tested still have a 50/50 chance of being a dud or live. This is much like the idea of flipping a coin. The shells do one of two things, explode or don't explode; just as flipping a coin has a result of heads or tails.

Starting with a big lot of 400 shells, let's assume that 50 percent of the shells are duds. One would think that if half the shells are duds, the testing would more often than not refuse the shipment of shells. What is the probability that a full lot of 399 or 398 shells would be sent to the battlefield?

50% Dud Artillery Shells



Looking at the shells independently, each shell has a 50/50 chance of working. The first shell has a 50 percent chance of exploding, in other words its chance of exploding is ½ and if the shell does explode the entire big lot of 399 is approved. The probability of 399 shells being accepted is one-half. If the first shell was a dud then another shell is picked and it also has ½ a chance of exploding or being a dud. If the second shell does explode, then the remaining 398 shells are sent to the battle field. If the second shell is also a dud, the sub-lot of 98 is refused and they move on to the next sub-lot.

Looking at this case the chance of the first shell being live was 50% or (½). The probability that the second shell will explode is determined by the chance that the first shell was a dud (50%) and then second shell explodes (50%). The second shell's probability of exploding is found by multiplying the two events that must happen together (½ x ½). To find the probability of either 398 or 399 shells being accepted, we add the two probabilities together and get a solution.

Total Probability = P(399) + P(398)

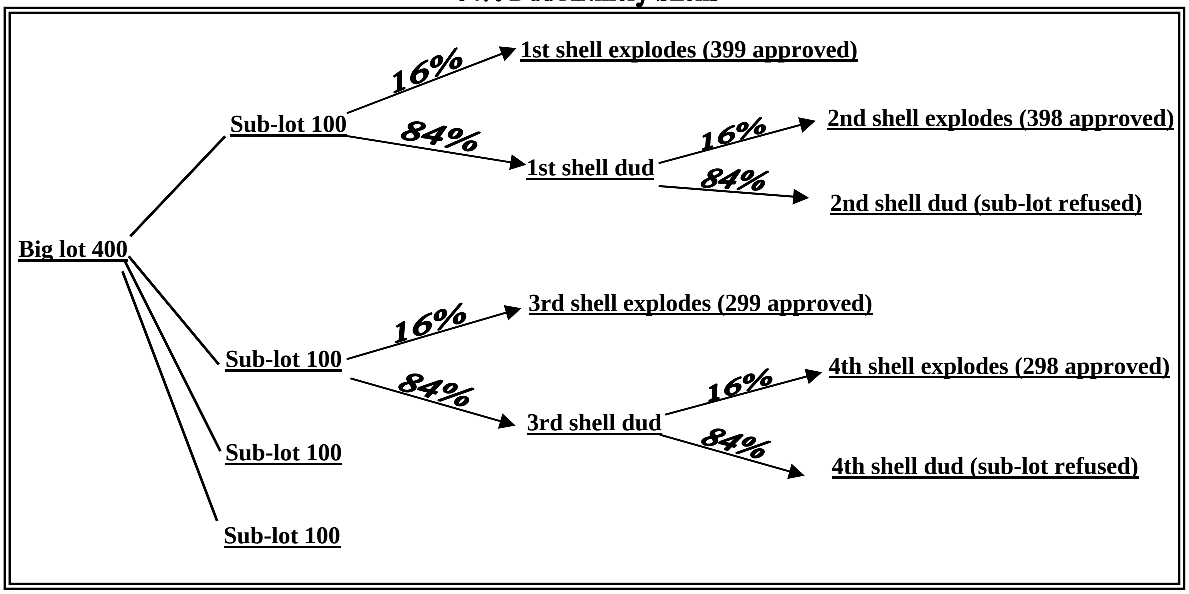
Total Probability = (½) + (½ · ½)

Total Probability = 0.75 = 75%

This means seventy-five percent of the time, when half of the shells were duds they would ship the remaining 398 or 399 shells into battle.

Taking a look at a second scenario we examine a higher percentage of failure. What if 84 percent of the shells were duds? We will find the probability of at least 298 of the shells being accepted and sent onto the battlefield for use. With 84 percent being duds that means only 16 percent of the shells were exploding. Using the same flow chart as earlier we can see how the testing progresses.

84% Dud Artillery Shells



Starting with the first sub-lot you see the first shell exploding would result in the 399 remaining shells being passed. The probability of 399 shells being accepted is (.16). If the first shell is a dud another shell would once again be chosen. The explosion of the second shell would leave 398 shells being approved. To find the probability of accepting 398 shells we multiply (.84) and (.16). However, if the second shell fails the entire first sub-lot is now removed. This would cause the remaining shell count to be at 300. The 300 shells still have the same independent percentage of failure and go through the same flow chart only with the exploding shells being numbered 299 and 298 respectively.

To find the probability of this problem we look at the path of the testing and the probability that leads to the final accepting of the lots. For 399 shells to be accepted it would be a probability of the first shell exploding 16% or (.16). To reach 398 accepted shells the path would flow to the dud part first (.84) and then to the exploding side (.16).

To arrive at the number of 299 shells there would have to be two duds (.84)(.84) and then one exploding shell (.16). Finally, for the 298 shells to be accepted would require three duds and the final shell exploding (.84)(.84)(.84)(.16). To find the total probability of 399, 398, 299, or 298 shells being accepted from the big lot, we must add all of these independent probabilities up.

Total Probability = P(399) + P(398) + P (299) + P(298)

Total Probability = (.16) + (.84)(.16) + (.84)(.84)(.16) + (.84)(.84)(.84)(.16)

Total Probability = $(.16) + (.84)(.16) + (.84)^2(.16) + (.84)^3(.16)$

Total Probability $\approx$ 0.5021

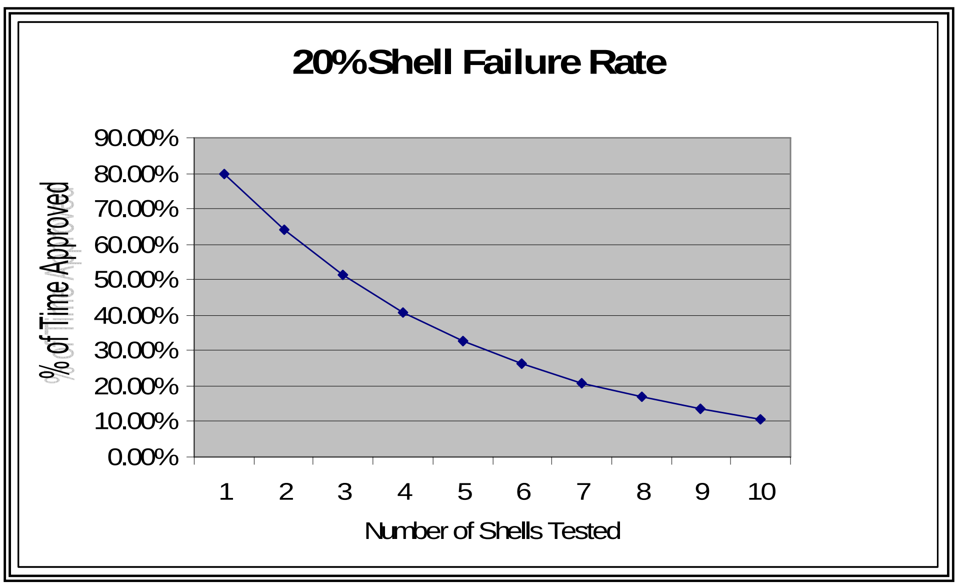
The 0.5021 represents slightly more then 50 percent of the time, when 84% of the shells were duds; they would send at least 298 shells into the battlefield. It may make one ask the question, was this process of testing adequate? I think this example would give a resounding no.

Perhaps a change in testing methods is needed. Let's instead start with a lot of 100 shells with a failure rate of 20 percent. This time we will test ten shells and if there are any duds found, the entire lot is rejected. We will ignore the possibilities that include dud shells, because any duds will result in the entire lot being rejected immediately. Instead we will focus on the probability that the shells will

100 Shells, 20% Failure Rate		
Shells Tested	**Probability**	**Percent of Time Approved**
1	$(0.8)^1$	80.00%
2	$(0.8)^2$	64.00%
3	$(0.8)^3$	51.20%
4	$(0.8)^4$	40.96%
5	$(0.8)^5$	32.77%
6	$(0.8)^6$	26.21%
7	$(0.8)^7$	20.97%
8	$(0.8)^8$	16.78%
9	$(0.8)^9$	13.42%
10	$(0.8)^{10}$	10.74%

explode. The table on the previous page shows the probability that the lot would be approved as the shells are tested, up to the required ten shells.

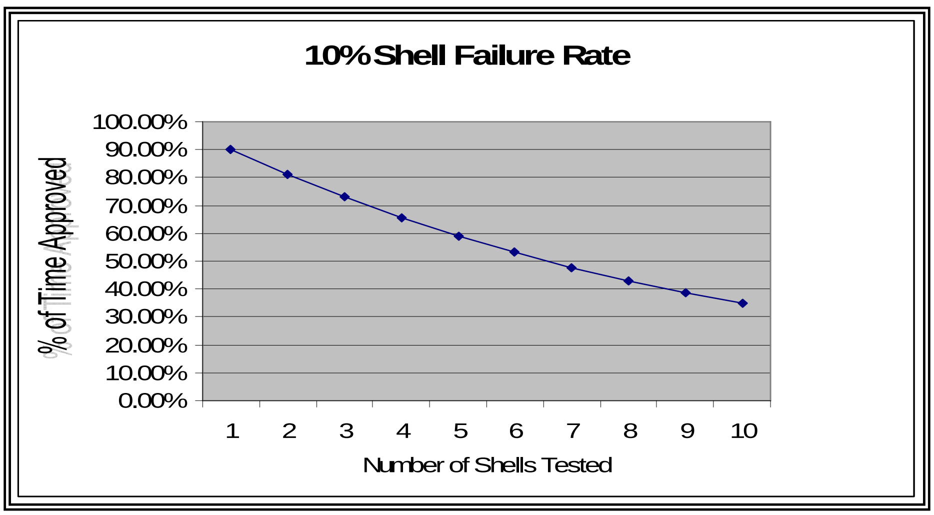
As shown, to find the total probability we take the probability of a shell exploding and raise it to the power of the shell number being tested. For example, to find the probability of accepting the lot testing ten shells the previous nine shells must all explode, each having an 80% probability of exploding. This leaves ninety shells to send on to war: $P(90) = (0.8)^{10}$. Taking data from the chart and making a graph helps show how the probability of the lot being approved decreases with each shell being tested.



They would accept a lot of shells that has a 20 percent failure rate just 11% of the time. Looking at the chart may help give us a better understanding of how many shells we could test to make sure the percentage of duds is low but the approval rating is high. As shown by the chart, as they continue to increase the number of shells tested, the probability of the lot being approved continues to decrease.

Altering the failure rate we will look at a test with only 10 percent of the shells being duds. Requiring that ten shells must all pass the testing for the lot to be approved we find the answer and use a more efficient way of finding the solution. To find the probability of accepting the lot of shells after testing ten out of the lot of one-hundred we take the probability of a shell exploding and raise it to the power of the number of shells we are testing. As shown by the chart and the graph, this results in the approval of the shell lot 35% of the time and rejecting of the shell lot 65% of the time.

100 Shells, 10% Failure Rate		
Shells Tested	Probability	Percent of Time Approved
1	$(0.9)^1$	90.00%
2	$(0.9)^2$	81.00%
3	$(0.9)^3$	72.90%
4	$(0.9)^4$	65.61%
5	$(0.9)^5$	59.05%
6	$(0.9)^6$	53.14%
7	$(0.9)^7$	47.83%
8	$(0.9)^8$	43.05%
9	$(0.9)^9$	38.74%
10	$(0.9)^{10}$	34.87%



The 35% can be misleading however, because it is counting the ten shells that were blown up as part of the lot sent on to the battlefield. If we want to compute the probability that a shell is sent to the field then we must find 90% of $(0.9)^{10}$ to get the probability that 90 remaining shells that are being sent on; (.90)(.35) = 0.315. This means that merely 31.5% of the shells were actually making it to the field when only 10% of the shells were duds.

Is there a better way to test for "proof"? When looking for answers we have to consider the problem. With objects like artillery shells, once you are done testing them, they cannot be used again. Other objects that would have the same type of "use it once" testing would include fireworks and crash testing vehicles for safety. You cannot test each and every object to see if they are working properly because you would have nothing left to use when you were done. Instead a better method of "proof" should be developed.

What is needed is a fast, yet efficient, way of figuring the probability; a way to test for proof and to help look for improved artillery shell sampling techniques. Then a case by case comparison could be done. The Binomial Distribution Probability Equation is one efficient way to estimate probability for our situation.

The Binomial Distribution Probability Equation

$$\text{Probability} = [nCx \cdot p^{x} \cdot q^{(n-x)}]$$

n = number of trials
x = number of successes among n trials
p = probability of success in any one trial
q = probability of failure in any one trial (*q* = 1 - *p*)
nCx = combinations of *n* items, choose *x*

Below are the requirements for using the Binomial Distribution Probability Equation; you can see that some of the cases we have done previously follow these requirements.

- Each trial can have only two outcomes and can be considered as either success or failure.
- There must be a fixed number of trials.
- The outcomes of each trial must be independent of each other.

Using the equation we try a few new trials of testing the artillery shells. If we set a minimum of acceptance rate as being 20% duds and we test 10 shells out of 100, we would expect that two shells would be duds. When using the Binomial Distribution Probability Equation for approximation we should make sure that our number of trials (*n*) compared to our lot size (100) is acceptable. In this case 10/100 = 0.1, (or the lot size is at least ten times the sample size); this is an acceptable range and we can proceed with using the formula.

With this in mind let's first examine accepting only one failure out of the ten tested shells. This means that when choosing ten random shells that the entire remaining lot of 90 will be approved if one shell is a dud and if no shells are duds. We can approximate the probability by using the following binomial probability formula. Converting this into the formula:

Probability $= [nCx \cdot p^x \cdot q^{(n-x)}]$

Probability = P(1 dud) + P(0 duds)

Probability $= 10C9 \cdot (.8)^9 \cdot (.2)^1 + 10C10 \cdot (.8)^{10} \cdot (.2)^0$

Probability $\approx$ 0.376 or 38%

The lot of artillery shells would be accepted 38% of the time. With the allowance of one dud in our test we raised the acceptance rate from 11% (on a previous problem) to 38% on this problem.

Using the same number of shells and same percentage of duds, but this time accepting two duds out of the ten shells we try the formula again.

$$\text{Probability} = [nCx \cdot p^{x} \cdot q^{(n-x)}]$$

Probability = P(2 duds) + P(1 dud) + P(0 duds)

$$\text{Probability} = 10C8 \cdot (.8)^{8} \cdot (.2)^{2} + 10C9 \cdot (.8)^{9} \cdot (.2)^{1} + 10C10 \cdot (.8)^{10} \cdot (.2)^{0}$$

$$\approx 0.678 \text{ or } 68\%$$

As shown the minor change of allowing a few duds makes a huge difference in the accepting of a perfectly good shell lot. The lot approval rate has now gone up to 68% by just allowing two duds out of ten.

Will this same line of testing make sure that lots with large numbers of duds are not approved? Letting our lot be 50% duds and testing it with the same allowable "two duds out of ten" rule we look at the formula.

$$\text{Probability} = [nCx \cdot p^{x} \cdot q^{(n-x)}]$$

Probability = P(2 duds) + P(1 dud) + P(0 duds)

$$\text{Probability} = 10C8 \cdot (.5)^{8} \cdot (.5)^{2} + 10C9 \cdot (.5)^{9} \cdot (.5)^{1} + 10C10 \cdot (.5)^{10} \cdot (.5)^{0}$$

$$\approx 0.055 \text{ or } 5.5\%$$

This is exactly what we would hope to happen. The lot with a high percentage of dud shells is refused a greater amount of time and the lot with a low percentage of dud shells is approved a greater amount of time.

Why could this method perhaps be better than the method used in WWII by the Ordnance Board? It appears their method was an “all or nothing” approach. If one artillery shell was good, then they all must be good and if two are bad then they all must be bad. Perhaps they also felt they were wasting too many shells if they tested up to ten shells out of each one-hundred lots. The bigger waste seems to be the amount of duds that were sent into battle.

As we can see, the testing techniques that the British Ordnance Board used in WWII were less then adequate. There are various techniques that could improve the “proof” testing including varying the number of shells being tested, altering the acceptance number of dud shells, and looking at more efficient ways of figuring the percentages. With increased testing and studying of data, one can better understand the needs and the desires of outcomes to help find the results that are satisfactory for “proof testing”.