

# Exact Sample Size for Special Nuclear Material Inventory Verification

James G. Wendelberger

Los Alamos National Laboratory, Mailstop: GS735, Los Alamos, NM 87544

## Abstract

The tables to compute the exact sample sizes for special nuclear material are provided based upon United States Department of Energy (DOE) sampling parameters for physical inventories for a population size up to 15,000 items. The paper motivates the exact computation of item sizes by showing exact correspondence to integer number of defects and enforcing the constraint of exactly attaining the minimum detectable defect rate. This is accomplished with a bivariate number of defects population mixture distribution. The number of defects in the bivariate mixture being sequential and bracketing the possibly non-integer defect number obtained by multiplying the minimum detectable defect rate times the population size (expected number of defects). There are allowances described and made for populations which would not have at least one defect based upon the minimum detectable defect rate. R code and output is provided in easy to follow examples.

LANL report numbers: LA-UR-15-23594 (text) and LA-UR-15-20692 (abstract) and LA-UR-15-25240 (presentation).

**Key Words:** Sample Size, Population, SNM, Mixture, Inventory

## 1. Overview

Los Alamos National Laboratory (LANL) conducts inventories of 100% of the items in each material balance area (MBA) except for the vault at Technical Area (TA)-55. Those items will be verified as being present and the appropriate tamper indicating device (TID) is applied and is intact. Items on inventory must be based on measured values. Items without TIDs are subject to a verification measurement dependent on a statistical selection. LANL has implemented a bi-monthly inventory for its Category I and II processing MBAs. Certain Category I MBAs have deviations that allow extended inventory periods. When the inventory is scheduled to commence, affected MBAs are locked in Local Area Nuclear Material Accountability Software (LANMAS) so that no transactions can be entered. A statistical verification sampling program generates a listing of items for verification measurement. A listing of the items is provided to the inventory personnel and contains item location, item identification and TID identification. A graded approach based on an item's characteristics is used to determine what attributes must be verified.

This document is based in part on the document by the similar name but revised March 2013, see [1].

## 2. Sampling Methodology

### 2.1 Defining the Population

The sample population for verification measurements includes:

- Materials without Tamper Indicating Devices
- Containers with Materials without Tamper Indicating Devices
- Category B and C Summary Material Types (20,40,44,45,50,70,82) materials  
>= 50 grams

A single population for verification measurements at TA-55 combines all items from TA-55 MBAs subject to inventory for the Physical Inventory Period except for the items in the vault.

A single population for verification measurements at CMR combines all items from all CMR MBAs for the Physical Inventory Period.

For MBAs outside of the TA-55 and CMR facilities a separate inventory verification sample is pulled for each outside MBA.

The sample population is stratified by category and the number of randomly selected items is based on the population of each category to meet the statistical Confidence Levels and minimum defects to meet Federal requirements, such as, 6.5.1.3 of [2]. Example: if a population of Category I items is 100, then 63 items must be verified to meet the 95% Confidence Levels and 3% Detectable Defects, see Table 2.

### 2.2 Defining Sampling Parameters

The sampling objectives are satisfied by applying a statistical statement about the entire Physical Inventory population. Assuming that “0” defects are identified during the Physical Inventory, the minimum sample size selected from each physical inventory population (Tables 2, 3 and 4) allows LANL to make a statistical statement that at a 95% confidence level, at least 97% of the population is “good” (3% minimum detectable defects) for Category I, at least 95% of the population is “good” (5% minimum detectable defects) for Category II and at least 90% of the population is “good” (10% minimum detectable defects) for Category III & IV item populations. These statements correspond to each of the different category levels as outlined in DOE-STD-1194-2011 [2], Table 6.5-1, Minimum Sampling Parameters for Physical Inventories, for the sampling parameters as outlined in Table 1 below.

**Table 1:** Minimum Sampling Parameters for Verification/Confirmation Measurements

Category	Confidence Level	Minimum Detectable Defects
I	95%	3%
II	95%	5%
III & IV	95%	10%

### 2.3 Sampling Algorithm

The design of the sampling plan is based on methods from Bowen and Bennett [3]. A pure hypergeometric distribution or mixture of two pure hypergeometric distributions will be the assumed model and the sample size will be calculated based off of the model.

Define the following notation:

p = probability of observing exactly zero defectives in the sample  
 n = sample size  
 N = finite population size  
 d = number of defectives  
 ! = factorial operator

Using the pure hypergeometric distribution, the probability of observing zero defects (denoted as p) is:

$$p = \frac{(N - d)! (N - n)!}{N! (N - d - n)!}$$

This pure hypergeometric probability formula or equation relates the sample size with a given confidence,  $1 - p$ , that a defective will be detected. Though the sample size,  $n$ , cannot be solved for directly in the above equation, numerical methods are employed to find the value of  $n$  that satisfies this equality. For a description of the algorithm used, see the Appendix 1: Numerical Algorithm for Finding a Sample Size at the end of this document.

The required sample sizes for population sizes from 1 to 15,000 for a given category level inventory population are given in Tables 2, 3 and 4 respectively. For a population size larger than 15,000 one may use the R code in Appendix 1: R Code to Create Sample Size Tables, to determine a sample size. It is likely to be nearly equal to the sample size listed for population size 15,000. This is because at some point the finite sample size is large enough to be considered infinite. In this case, the sample size is fixed to some value, a value that is likely near the largest sample size value in the correct table. For population sizes in the range 15,000 to 10,000,000 the sample size needed is the same as that for 15,000 for all three category levels (99, 59 and 29). This is checked in Mathematica with the code in Appendix 3: Mathematica Code to Extend the Sample Size Tables to Population Size 10,000,000. This effectively extends the Tables 2, 3 and 4 to populations as large as 10,000,000.

The output of this code, see Appendix 3 Output section, reveals that the probability, attained significance level, for the smaller sample size is not adequate (it is greater than .05) and that for a sample size one larger the sample size is adequate (attained alpha is less than .05). The number of defects in this case is an integer so the pure hypergeometric distribution probability equation in the Sampling Algorithm section is used directly. The number of defects is an integer so there is no need to use the mixture distribution. The smaller sample size produces a probability greater than .05 and the larger sample size (by one item) produces a probability less than .05 for each of the three category levels.

How do we handle a non-integer number of defects? One could simply evaluate the pure hypergeometric distribution probability with a non-integer number of defects. This is possible and requires one to generalize the factorial in the formula to non-integer values. Mathematicians have done this using the Gamma function. However, that methodology, simply generalizing the evaluation of the pure hypergeometric probability formula for  $p$ , by using the non-integer value of  $r$  in the formula, as we will see below may not produce the correct sample size. Instead we will employ a mixture distribution methodology when the number of defects is a non-integer. This mixture methodology brackets the non-integer number of defects by nearest smaller and larger integers and then uses these integer numbers in a mixture of two pure hypergeometric distributions as described in the appendix Numerical Algorithm for Finding a Sample Size to find the appropriate sample size. This way we never need to use the hypergeometric probability formula on a non-integer number of defects. In fact we will also see that when we have a number of defects less than one then we will instead use one defect, an integer number, to solve both this non-integer problem and another problem to be described later.

Let us look at an example. Consider the case  $N = 78$  and  $r = .03$ . Then  $d = N \times R = 78 \times .03 = 2.34$ . Notice that if we incorrectly use the pure hypergeometric probability formula directly in Mathematica we obtain a sample size of size 56, see Appendix 4. This is because Mathematica generalizes the factorial to the Gamma function which does not hold here. As previously noted we cannot have 2.34 defects. Instead we use the mixture distribution of two hypergeometric distributions to solve this problem. Form the mixture distribution with 66% having 2 defects and 34% having 3 defects. This still gives the correct expected number of defects as:  $2 \times .66 + 3 \times .34 = 1.32 + 1.02 = 2.34$  and defect rate = .03 for a population of size 78. However, now instead of finding the sample size which gives 95% confidence and a non-integer defect number we determine the sample size for the mixture distribution which uses integer numbers of defects. In this case we want the sample size which provides less than 5% probability of 2 defects 66% of the time and 3 defects 34% of the time. The Mathematica code to compute this (correctly) is provided in Appendix 4.

The mixture distribution properly proportions the two pure hypergeometric distributions, using an integer number of defects each, to give the correct assumed defect rate of 3%. The output of this is  $\{0.0520979, 0.0468531\}$  showing that a sample size of 57 gives less than 95% confidence ( $1 - 0.0520979$ ) and a sample size of 58 gives greater than 95% confidence ( $1 - 0.0468531$ ). Thus, we have found that the correct sample size for  $N=78$ ,  $r=.03$  and 95% confidence is 58. This is larger than the direct use of the non-integer value in the pure hypergeometric probability formula which we calculated is 56. The direct use of the hypergeometric probability formula with non-integer defect number values yields the wrong answer, 56 instead of 58. This is the reason to use the mixture distribution to provide the correct sample size for the assumed defect rate(s).

This example shows that the direct use of the non-integer defect number in the probability formula yields a sample size that is too small. This should not be a surprise as the probability formula uses the hypergeometric distribution which assumes an integer

**Table 2: Sample Size Required for Cat I NM Inventory**

95% Confidence / 3% defects*	
Population (N)	Sample Size
1 through 19	N
20 through 36	N-1
37 through 43	N-2
44 through 47	N-3
48 through 51	N-4
52 through 53	N-5
54 through 55	N-6
56 through 57	N-7
57 through 58	50
59 through 62	51
63 through 67	52
68 through 69	53
70 through 71	54
72 through 73	55
74 through 75	56
76 through 77	57
78 through 79	58
80 through 81	59
82 through 84	60
85 through 87	61
88 through 92	62
93 through 101	63
102 through 103	64
104 through 106	65
107 through 109	66
110 through 113	67
114 through 118	68
119 through 124	69
125 through 135	70
136 through 139	71
140 through 143	72
144 through 149	73
150 through 159	74
160 through 170	75
171 through 177	76
178 through 185	77
186 through 202	78
203 through 210	79
211 through 224	80
225 through 240	81
241 through 256	82
257 through 276	83
277 through 302	84
303 through 323	85
324 through 351	86
352 through 386	87
387 through 434	88
435 through 478	89
479 through 542	90
543 through 619	91
620 through 725	92
726 through 871	93
872 through 1076	94
1077 through 1409	95
1410 through 2031	96
2032 through 3553	97
3554 through 13761	98
13762 through 15000	99

\* - At least one defect in the population

**Table 3: Sample Size Required for Cat II NM Inventory**

95% Confidence / 5% Defects*	
Population (N)	Sample Size
1 through 19	N
20 through 26	N-1
27 through 30	N-2
30 through 31	28
32 through 33	29
34 through 35	30
36 through 40	31
41 through 42	32
43 through 44	33
45 through 46	34
47 through 48	35
49 through 51	36
52 through 55	37
56 through 61	38
62 through 64	39
65 through 68	40
69 through 73	41
74 through 82	42
83 through 87	43
88 through 94	44
95 through 104	45
105 through 113	46
114 through 126	47
127 through 142	48
143 through 161	49
162 through 182	50
183 through 207	51
208 through 244	52
245 through 293	53
294 through 366	54
367 through 482	55
483 through 690	56
691 through 1202	57
1203 through 4233	58
4234 through 15000	59

\* - At least one defect in the population

**Table 4: Sample Size Required for Cat III & IV NM Inventory**

95% Confidence / 10% Defects*	
Population (N)	Sample Size
1 through 13	N
14 through 16	N-1
16 through 18	15
19 through 21	16
22 through 23	17
24 through 26	18
27 through 31	19
32 through 36	20
37 through 43	21
44 through 52	22
53 through 63	23
64 through 81	24
82 through 108	25
109 through 156	26
157 through 274	27
275 through 937	28
938 through 15000	29

\* - At least one defect in population

value of the number of defects. However, the computer code to compute the factorial will do so for a non-integer value as well. What we have shown is that the improper use of this formula for non-integer defect numbers may yield an incorrect sample size for the problem solution.

As mentioned earlier the Tables 2, 3 and 4 were computed using the R code in Appendix 2. The Mathematica code in Appendix 5: Mathematica Code to Replicate/Verify Tables Created by R Code was used to check that this code gave the same result for Table 4 and for the first 100 population sizes of Tables 2 and 3. This verification means we have computed the sample sizes with two different codes on two different software packages, R and Mathematica, and, we obtain the same resulting sample sizes.

The mixture formula only applies when the expected number of defects is greater than or equal to one. This issue was alluded to previously. For this reason and for the peculiar nature of assuming a defect rate that is smaller than the alpha level we choose to work with one defect in the population when the defect rate times the population would yield less than one defect. This has the benefit of being calculable by the pure hypergeometric probability formula and in the case where the defect rate is less than or equal to the attained alpha level (1 - confidence level) it removes the peculiar case when  $N$  is equal to one and it increases by one the similarly peculiar cases when  $N$  is greater than one. This choice of how to handle defect rates which provide a value of less than one defect is thereby conservative. It provides a census (the entire population) as the sample for all population sizes which have an expected number of defects less than one.

The sample items are identified using the procedures described in S4-DTG-060, "Generation of the Physical Inventory Sample". The inventory process is carried out according to S4-DTG-003, "Physical Inventory."

### 3. Appendices

#### 3.1 Appendix 1: Numerical Algorithm for Finding a Sample Size

This appendix describes the algorithm used for the selection of the sample size needed for the design of the sampling plan. Let  $X$  denotes the random variable associated with the number of defects in a random sample of size  $n$  drawn from a population of size  $N$  with  $d$  defective items. Then  $X$  follows a hypergeometric distribution. Assume that the population size ( $N$ ) is fixed and assume a value of the proportion of defects in the population ( $r$ ). In order to use the equation in the mathematics section, we need  $d$ , the number of defective items in the population. This value is calculated as  $d = N \times r$ . Notice, however, that this value may not always be integer valued.

The direct use of the pure hypergeometric distribution requires an integer number of defects in the population. The following example illustrates the methodology. In this example, the population size is 90,  $N = 90$ . The prescribed defect rate,  $r$ , is 3%. This results in  $90 \times 3\%$  or 2.7 defects in the population. In the real world there cannot be a fractional number of defects. An item is either defective or not. Correspondingly, in order to utilize the pure hypergeometric probability function the number of defects must also be

an integer. To accommodate this integer restriction we utilize a mixture distribution. This mixture distribution has a probability of a population with 2 defects and the remaining probability of a population with 3 defects. The fraction of all items in the mixture population with 2 defects is 30% and that with 3 defects is 70%. These numbers are chosen so that the expected number of defects in the mixture distribution is 2.7. This is accomplished mathematically by solving the equation for the fraction,  $f \times 2 + (1 - f) \times 3 = 2.7$  yielding  $f = 30\%$ . In summary, we may think of this as having a population distribution, each with a total of 90 items, 30% of which have 2 defects and 70% of which have 3 defects. This provides a distribution which we may utilize to determine the required sample size to be 95% confident of less than 3% defects. Call this mixture distribution M. The mixture distribution M matches the one when  $N \times r$  is an integer and provides a monotonic path from integer value to integer value for the integer number of defects.

M is a distribution of 90 items 30% of which have 2 defects and 70% of which have 3 defects. If we choose a sample of size  $s$  from the distribution M then we want the probability that we observe zero defects at random or by chance to be less than 5% (100% - 95%). This means that when we actually observe zero defects in a sample of size  $s$  that we can be 95% confident that the defect rate is less than 3%. Writing this mathematically we want the smallest sample size  $s$  so that the Probability( zero defects | random sample of size  $s$  from M) to be at most 5%. The random sample is chosen without replacement. Using the fact that M is a mixture distribution find the smallest  $s$  so that

$$.3 \times \text{Probability( zero defects | random sample of size } s \text{ from a size 90 distribution with 2 defects)} +$$

$$.7 \times \text{Probability( zero defects | random sample of size } s \text{ from a size 90 distribution with 3 defects)}$$

$$\leq 5\%.$$

Both probabilities in this minimization formula are pure hypergeometric probabilities which can be computed using the pure hypergeometric probability equation provided in the Sampling Algorithm section. Using this pure hypergeometric probability equation we obtain:

$$.3 \frac{(90 - 2)! (90 - s)!}{(90)! (90 - 2 - s)!} + .7 \frac{(90 - 3)! (90 - s)!}{(90)! (90 - 3 - s)!} \leq 5\%$$

Running this calculation in R we calculate the left hand side for  $s = 62$  as

$$.3 * (\text{factorial}(90 - 2) * \text{factorial}(90 - s)) / (\text{factorial}(90) * \text{factorial}(90 - 2 - s)) + .7 * (\text{factorial}(90 - 3) * \text{factorial}(90 - s)) / (\text{factorial}(90) * \text{factorial}(90 - 3 - s))$$

equal to 0.04783453 which is less than 5%. Then running this calculation in R we calculate the left hand side, for  $s = 61$ , equal to 0.0521842 which exceeds 5%. Thus the sample size for  $N=90$ , defect rate = 3% at 95% confidence is 62.

The methodology used in this example may be used for any defect rate that has a mixture distribution with at least one defect. For those with less than one defect there is a modification because the Probability( zero defects | random sample of size  $s$  from a size  $N$  distribution with 0 defects) is identically equal to 1 for any sample size. In this case the overall probability is driven entirely by the second term in the sum, that is,  $(1 - f) \times$  Probability( zero defects | random sample of size  $s$  from a size  $N$  distribution with 1 defect). However, in what we need here we can ignore this issue because we are going to require at least one defect in the population.

Another complication appears. Notice that for  $N=1$ , defect rate = 3% and 95% confidence we would have a sample size of zero. This is because the defect rate is already less than 5% so if we assume the defect rate is 3% we do not need to sample to get 95% confidence. In fact we have 97% confidence. This is not a common sense thing to do. It is related to the fact that the minimum number of non-zero defects that we can have is in fact an integer and the integer one. In order to overcome this anomaly we will use the defect rate that corresponds to one defect whenever the assumed defect rate would imply a number of defects that is less than one. The result of this for the parameters needed here is to increase some of the lower population sample sizes by one. In other words it only results in a slightly larger sample (by one) while removing the above nonsensical result. For this reason we call the sample sizes computed in Tables 2-4 “adjusted”. The sample size is adjusted up by one for some of the smaller population sizes.

In summary the algorithm to compute the desired sample size  $S$  in general is as follows:

1. Determine the population size  $N$ , the defect rate  $r$  and the desired confidence level  $c$ .
2. Compute the expected number of defects =  $N \times r$ .
3. Determine the mixture distribution  $M$ . If the expected number of defects is less than one then set the lower number of defects,  $L$ , equal to one, the upper number of defects,  $U$ , equal to 2 and the fraction  $f = 1.0$  otherwise set the lower number of defects  $L = \text{floor}(N \times r)$ , the upper number of defects  $U = \text{floor}(N \times r) + 1$  and solve  $f \times L + (1-f) \times U = N \times r$  for  $f$ .
4. For  $s = 1, \dots, N$  compute

$$\alpha(s) = f \frac{(N - L)! (N - s)!}{(N)! (N - L - s)!} + (1 - f) \frac{(N - U)! (N - s)!}{(N)! (N - U - s)!}$$

5. Choose the sample size  $S$  as the smallest  $s$  for which  $\alpha(s) \leq 1 - c$ .

Here  $\text{floor}()$  is the function that truncates the decimal part of the real number. In step (3) when the expected number of defects is one we actually solve the problem with a different defect rate than the prescribed one of  $r$ . In this case, the new defect rate turns out to be  $r^* = 1/N$ . Thus,  $r^* \times N = 1$ , giving one defect in the population of size  $N$ .

In step 4 the quantity  $\alpha(s)$  is monotonically decreasing. This means that we only need to find the cutoff point or sample size  $s$  where  $\alpha(s - 1) > 1 - c$  and  $\alpha(s) < 1 - c$  to



determine the  $s$  value which gives a confidence of size  $c$ . In practice one may not need to compute  $\alpha(s)$  for all  $N+1$  values of  $s$  ( $0, 1, 2, \dots, N$ ). For example, one could start from the smallest (or largest)  $s$  and increase (or decrease)  $s$  by 1 until the cutoff point is determined.

### 3.2 Appendix 2: R Code to Create Sample Size Tables

This appendix contains the R code for the algorithm used for the selection of the sample size needed for the design of the sampling plan, in particular for Tables 2, 3 and 4.

```
createadjustedtable <- function(desiredconfidence,mdd,startpop,endpop) {
pop <- startpop
originalmdd <- mdd
myfile <-
paste(getwd(),"/Conf",desiredconfidence*100,"Defect",originalmdd*100,"Initial",startpop,
"End",endpop,".txt",sep="")
sink(myfile,append=FALSE,split=TRUE)
cat("Confidence"," ", "Defect Rate"," ", "Population ", "Sample Size ", "Attained
Confidence","\n")
sink()
while (pop <= endpop & originalmdd*pop < 1) {
mdd <- 1/pop
numdefmin <- 0
numdefmax <- numdefmin+1
numgoodmin <- pop-numdefmin
numgoodmax <- pop-numdefmax
largerfraction <- 1-numdefmin
smallerfraction <- 1 - largerfraction
testsamplesize <- 0
maxtestsamplesize <- pop
starttestsamplesize <- testsamplesize
samplesize <- 0
attainedconf <- 0.0
for (i in 0:maxtestsamplesize) {
if (numdefmin == 0) {attainedconf <- smallerfraction*
{phyper(0,numdefmin,numgoodmin,i)} + largerfraction*(1-
phyper(0,numdefmax,numgoodmax,i))}
if (numdefmin != 0) {attainedconf <- smallerfraction* {1-
phyper(0,numdefmin,numgoodmin,i)} + largerfraction*(1-
phyper(0,numdefmax,numgoodmax,i))}
if(attainedconf >= desiredconfidence) {samplesize<-i}
if(attainedconf >= desiredconfidence) {break}
}
sink(myfile,append=TRUE,split=TRUE)
cat(" ",desiredconfidence," ",formatC( round( mdd, 3 ), format='f', digits=3 ),"
",formatC( round( pop, 0 ), format='f', width=6,digits=0 )," ",formatC( round(
samplesize, 0 ), format='f', width=6,digits=0 )," ",formatC( round( attainedconf, 3
), format='f', digits=3 ),"\n")
sink()
pop <- pop + 1
}
```

```

mdd <- originalmdd
while (pop <= endpop & originalmdd*pop >= 1) {
numdefmin <- floor(mdd*pop)
numdefmax <- numdefmin+1
numgoodmin <- pop-numdefmin
numgoodmax <- pop-numdefmax
largerfraction <- mdd*pop-numdefmin
smallerfraction <- 1 - largerfraction
testsamplesize <- 0
maxtestsamplesize <- pop
starttestsamplesize <- testsamplesize
samplesize <- 0
attainedconf <- 0.0
for (i in 0:maxtestsamplesize) {
  if (numdefmin == 0) {attainedconf <- smallerfraction*
{phyper(0,numdefmin,numgoodmin,i)} + largerfraction*(1-
phyper(0,numdefmax,numgoodmax,i))}
  if (numdefmin != 0) {attainedconf <- smallerfraction* {1-
phyper(0,numdefmin,numgoodmin,i)} + largerfraction*(1-
phyper(0,numdefmax,numgoodmax,i))}
  if(attainedconf >= desiredconfidence) {samplesize<-i}
  if(attainedconf >= desiredconfidence) {break}
}
sink(myfile,append=TRUE,split=TRUE)
cat(" ",desiredconfidence," ",formatC( round( mdd, 3 ), format='f', digits=3 ),"
",formatC( round( pop, 0 ), format='f', width=6,digits=0 )," ",formatC( round(
samplesize, 0 ), format='f', width=6,digits=0 )," ",formatC( round( attainedconf, 3
), format='f', digits=3 ),"\n")
sink()
pop <- pop + 1
}
}

```

### 3.3 Appendix 3: Mathematica Code to Extend the Sample Size Tables to Population Size 10,000,000

Input:

```
NN:=10000000
```

```
d:=NN*.03
```

```
p:=((NN-d)!*(NN-n)!)/((NN)!*(NN-d-n!))
```

```
Table[p,{n,98,99}]
```

```
d:=NN*.05
```

```
p:=((NN-d)!*(NN-n)!)/((NN)!*(NN-d-n!))
```

```
Table[p,{n,58,59}]
```

```
d:=NN*.10
```

```
p:=((NN-d)!*(NN-n)!)/((NN)!*(NN-d-n!))
```

Table[p, {n,28,29}]

Output:

{0.05053864,0.04902247}

{0.05053864,0.04902247}

{0.05233454,0.04710107}

### 3.4 Appendix 4: Mathematica Code for Example with Population Size 78, 3% Reject Rate and 95% Confidence

Mathematica code using the incorrect formula for a fractional number of rejects.

Input:

NN:=78

d:=NN\*.03

p:=((NN-d)!\*(NN-n)!)/((NN)!\*(NN-d-n)!)

Table[p, {n,55,56}]

Output:

{0.0546156, 0.0490591}

Mathematica code using the correct formula for a fractional number of rejects.

1. Input:

2. NN:=78

3. d:=NN\*.03

4. d1:=Floor[d]

5. d2:=d1+1

6. f2:=d-d1

7. f1:=1-f2

8. p:=f1\*(((NN-d1)!\*(NN-n)!)/((NN)!\*(NN-d1-n)!))+f2\*(((NN-d2)!\*(NN-n)!)/((NN)!\*(NN-d2-n)!))

Table[p, {n,57,58}]

Output:

{0.0520979, 0.0468531}

### 3.5 Appendix 5: Mathematica Code to Replicate/Verify Tables Created by R Code

This code may be used to verify Table 4 except for the largest sample size which may be verified using Appendix 3 code.

Input:

```
f[NN_] := Module[{conf=.95,dr=.10,dstart},{alpha:=1-conf;dstart:=NN*dr ;
d:=Max[dstart,1];d1:=Floor[d];
d2:= d1+1;f2:=d-d1;f1:=1-f2;
p:=f1*( ((NN-d1)!*(NN-n)!) / ((NN)!*(NN-d1-n)!)+f2*(((NN-d2)!*(NN-
n)!)/((NN)!*(NN-d2-n)!))};
mylist:=Table[{p,n,NN,dr,conf},{n,0,NN},{NN}];
mylist2:=First[Select[Flatten[mylist,1],#[[1]]<=alpha& ]];
Print[mylist2]]
Do[f[pop],{pop,2,938}]
```

Output:

A table containing attained alpha, sample size, population, reject rate, and confidence level.

## Acknowledgements

Disclaimer:

Los Alamos National Security, LLC for the National Nuclear Security Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396. By approving this article, the publisher recognizes that the U.S. Government retains nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

LANL report numbers: LA-UR-15-23594 (text) and LA-UR-15-20692 (abstract) and LA-UR-15-25240 (presentation).

## References

1. SAFE-4 Statistical Inventory / Verification Measurements Sampling Plan: Revised March 2013.
2. DOE Standard: Nuclear Materials Control and Accountability, DOE-STD-1194-2011, Change Notice No. 2, December 2012.
3. W. M. Bowen and C. A. Bennett, editors, Statistical Methods for Nuclear Material Management, U.S. Nuclear Regulatory Commission, NUREG/CR-4604, U.S. Government Printing Office, Washington, D.C., 1988.