SANDIA REPORT

SAND2010-0550
Unlimited Release
Printed February 2010

Sample Sizes for Confidence Limits for Reliability

John L. Darby

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of Energy's
National Nuclear Security Administration under Contract DE-AC04-94AL85000.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Available to DOE and DOE contractors from

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone:	(865) 576-8401
Facsimile:	(865) 576-5728
E-Mail:	reports@adonis.osti.gov
Online ordering:	http://www.osti.gov/bridge

Available to the public from

U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd.
Springfield, VA 22161

Telephone:	(800) 553-6847
Facsimile:	(703) 605-6900
E-Mail:	orders@ntis.fedworld.gov
Online order:	http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2010-0550
Unlimited Release
Printed February 2010

Sample Sizes for Confidence Limits For Reliability

John L. Darby
Systems Analysis I
Sandia National Laboratories
P.O. Box 5800
Albuquerque, New Mexico 87185-MS0415

Abstract

We recently performed an evaluation of the implications of a reduced stockpile of nuclear weapons for surveillance to support estimates of reliability. We found that one technique developed at Sandia National Laboratories (SNL) under-estimates the required sample size for systems-level testing. For a large population the discrepancy is not important, but for a small population it is important. We found that another technique used by SNL provides the correct required sample size.

For systems-level testing of nuclear weapons, samples are selected without replacement, and the hypergeometric probability distribution applies. Both of the SNL techniques focus on samples without defects from sampling without replacement. We generalized the second SNL technique to cases with defects in the sample.

We created a computer program in Mathematica to automate the calculation of confidence for reliability. We also evaluated sampling with replacement where the binomial probability distribution applies.

ACKNOWLEDGMENTS

The author appreciated the careful reviews of this report by Douglas Loescher, Kathleen Diegert, and Steven Hatch of Sandia National Laboratories. Loescher provided the insight that for sampling without replacement, if the samples are never returned to the population, the reliability to a given confidence changes for the remaining population.

CONTENTS

1. INTRODUCTION 9
2. SAMPLING PROCESS 9
3. CONFIDENCE LIMITS 10
4. CONFIDENCE LIMITS FOR THE BINOMIAL DISTRIBUTION 10
5. CONFIDENCE LIMITS FOR THE HYPERGEOMETRIC DISTRIBUTION 12
6. SAMPLE SIZES FOR NUCLEAR WEAPONS RELIABILITY 14
7. CONCLUSIONS AND RECOMMENDATIONS 16

References 17
Appendix A. Mathematica Code 19
Distribution 27

TABLES

Table 1. Required Sample Size for Various Population Sizes for 90% confidence for 90% Reliability With No failures in the Sample Sampling without Replacement 15

Table 2. Required Sample Size for Various Population Sizes for 90% confidence for 95% Reliability With No failures in the Sample Sampling without Replacement 15

(THIS PAGE INTENTIONALLY LEFT BLANK)

EXECUTIVE SUMMARY

We recently performed an evaluation of the implications of a reduced stockpile of nuclear weapons for surveillance to support estimates of reliability. [Stockpile Surveillance] As part of this effort, we looked at the number of samples required to support statements about confidence for reliability; we considered very small population sizes of weapons.

For systems-level testing of nuclear weapons, samples are selected without replacement, and the hypergeometric probability distribution applies.

We found that one technique developed at Sandia National Laboratories (SNL) under-estimates the required sample size for systems-level testing. For a large population the discrepancy is not important, but for a small population it is important. We found that another technique used by SNL provides the correct required sample size.

Both of the SNL techniques focus on samples without defects from sampling without replacement. We generalized the second SNL technique to cases with defects in the sample. We created a computer program in Mathematica to automate the calculation of confidence for reliability. [Mathematica] We also evaluated sampling with replacement where the binomial probability distribution applies. Both sampling without and with replacement are addressed in this report, and techniques for calculating confidence for reliability for both sampling strategies are implemented in the Mathematica program.

Previous discussions of the sample size summarize the required sample size for sampling without replacement, given no defects in the sample, as a function of population size. [Hahn] Typically, both 90% confidence for 90% reliability and 90% confidence for 95% reliability are considered. Some of these discussions used the incorrect technique. We evaluated the required sample size for both cases using the correct and the incorrect technique. The results are as follows.

Table E-1.
Required Sample Size for Various Population Sizes
for 90% confidence for 90% Reliability
With No failures in the Sample
Sampling without Replacement

Population Size	Required Sample Size from SNL Second Technique	Required Sample Size from SNL First Technique (INCORRECT)
10	9	7
20	14	11
30	16	13
40	17	14
50	18	16
70	19	17
90	20	18
120	20	19
150	21	20
250	21	21
275	22[a]	21
532 or greater	22	22[a]

[a]The sample size for sampling with replacement required by the binomial distribution is 22.

Table E-2.
Required Sample Size for Various Population Sizes
for 90% confidence for 95% Reliability
With No failures in the Sample
Sampling without Replacement

Population Size	Required Sample Size from SNL Second Technique	Required Sample Size from SNL First Technique (INCORRECT)
10	9	8
20	18	14
40	27	21
100	37	32
200	41	37
300	42	40
400	43	41
500	43	42
800	44	43
1000	44	44
1200	45[b]	44
2131 or greater	45	45[b]

[b]The sample size for sampling with replacement required by the binomial distribution is 45.

1. INTRODUCTION

This report summarizes techniques to calculate classical statistical confidence limits for reliability based on sample results. Two sampling distributions are discussed:

1. Sampling with replacement, the binomial distribution, and
2. Sampling without replacement, the hypergeometric distribution.

For sampling with replacement, three techniques are discussed, including a simple one limited to the special case of no failures in the sample.

For sampling without replacement, three techniques are discussed. Two of the techniques are provided in Sandia National Laboratories (SNL) reports for estimating the reliability of nuclear weapons; these two focus on samples with no defective items found in the sample. The first technique is not as accurate as the second as subsequently discussed; specifically, the first technique under-predicts the required sample size for a small population.

In this report, the second technique for sampling without replacement is expanded to a third technique to address cases where there are defective items found in the sample.

All the techniques are implemented in a Mathematica program, ConfidenceLimitsIterative.nb, written by the author. The program is provided in Appendix A with example calculations; results were verified by comparison with published values.

2. SAMPLING PROCESS

Sampling with replacement means that each item selected from a population is replaced and can possibly be selected again. Sampling without replacement means that once an item is selected it is not returned to the population.

The binomial distribution is used to model sampling with replacement. The binomial distribution has two parameters p and n, and is denoted here as BinDist(p, n). p is the probability that an item in the population is failed and n is the number of items in the sample. p must be a constant and can be any value in [0, 1]. The binomial distribution is independent of the population size N. For a small population with a large sample, $n > N$, this means that some items in the population will be sampled- with replacement- more than once. Although the binomial distribution is mathematically applicable for $n > N$, the extent to which a small sample is representative of the population is questionable.

The hypergeometric distribution is used to model sampling without replacement. The hypergeometric distribution has three parameters N, n, and D, and is here denoted as HyperDist(n, D, N). N is the population size, n is the sample size, and D is the number of defective items in the population. D must a be discrete value {0, 1, 2, … N}. The probability that an item in the population is defective is D/N and is restricted to {0, 1/N, 2/N, …, 1}.

The probability of selecting a defective item changes as the sample is taken, since once an item is selected it is not returned to the population.

Let R denote the reliability. For sampling with replacement R = 1 - p; for sampling without replacement, R = 1 – D/N.[1]

For simplicity of nomenclature, “probability of failure” denoted as P is defined to mean either p or D/N depending on the type of sampling being discussed, with or without replacement, respectively. R = 1 - P; R can be expressed as a percent.

3. CONFIDENCE LIMITS

P is not known but is estimated from the sample.[2] Due to uncertainty, P is expressed with an upper one-sided confidence limit.[3] Let the probability that P is in the interval $[0, UCL_\gamma]$ be $1-\gamma$. UCL_γ is called a $(1-\gamma)100\%$ upper one-sided confidence limit for P. [Martz] If P is in $[0, UCL_\gamma]$ with probability $1-\gamma$, R is in $[1 - UCL_\gamma, 1]$ with probability $1-\gamma$.

For simplicity of nomenclature, “confidence” is defined to mean $(1-\gamma)100\%$.

For example, for 50% confidence γ is 0.5; if $UCL_{0.5}$ is 0.3, the probability that P is in [0, 0.3] is 0.5.For 90% confidence γ is 0.10; if $UCL_{0.1}$ is 0.6, the probability that P is in [0, 0.6] is 0.9.

Statements about the confidence of reliability specify $1 - UCL_\gamma$. For example “90% confidence for 95% reliability” means $1 – UCL_{0.1}$ is 0.95.

In the following, $\mathbb{R}$ will be used to denote the specific value $1 - UCL_\gamma$, and $\mathbb{P}$ will denote the specific value UCL_γ.

4. CONFIDENCE LIMITS FOR THE BINOMIAL DISTRIBUTION

For sampling with replacement, UCL_γ depends on the confidence desired, the size of the sample, and the number of failures in the sample; it does not depend on the size of the population. Let $UCL_\gamma(x, n)$ denotes UCL_γ for confidence $(1-\gamma)100\%$ for a sample of size n with x failures.

---

[1] If the samples are never returned to the population (perhaps because they were destroyed), the reliability for the population of size N-n that is left changes as a result of the permanent removal of items. Suppose a sample of size 37 is drawn from a population of 100 and no failures are observed. It is correct to assert with 90% confidence that that the original population did not contain more than five defective units, 95% reliability. But we know that there were no defective units in the sample, so the five possible defective units could only be in the remaining population of 63 units. For the remaining population, the reliability at 90% confidence is decreased to a little more than 92%. [Loescher]

[2] Classical probability treats the probability as fixed but perhaps unknown. [Dougherty] A Bayesian approach considers probability as a random variable; Bayesian confidence limits are not addressed in this report.

[3] Two sided confidence limits can also be generated. [Martz]

For the case of no failures, SNL uses a simple formula to calculate $UCL_{0.5}(0, n)$: [SNL Weapon Reliability Guide]

$$UCL_{0.5}(0, n) = 1 - 0.5^{1/n} \qquad \text{(Eqn. 1)}$$

With no failures in the sample this simple formula can be generalized for any desired confidence:

$$UCL_{\gamma}(0, n) = 1 - \gamma^{1/n} \qquad \text{(Eqn. 2)}$$

Considering any number of failures in the sample: [Martz]

$$UCL_{\gamma}(x, n) = \frac{(x+1)F_{1-\gamma}(2x+2, 2n-2x)}{(n-x) + (x+1)F_{1-\gamma}(2x+2, 2n-2x)} \qquad \text{(Eqn. 3)}$$

where F is the F-ratio distribution. For no failures, equation 3 and equation 2 provide identical results.

A third technique for calculating $UCL_{\gamma}(x, n)$ directly from the Cumulative Distribution Function (CDF) is as follows. [Beyer, Section III.3] For the binomial distribution BinDist(p, n), find the p for which the CDF evaluated at x equals γ; that p is $UCL_{\gamma}(x, n)$. For example, for 95% confidence given x failures in n samples, the p for which CDF(BinDist(n, p), x) equals 0.05 is $UCL_{0.05}(x, n)$.

For the special case of no failures, CDF(BinDist(n, p), x) equals PDF(BinDist(n, p), 0) where PDF is the Probability Density Function.

The technique using the CDF provides results identical to those obtained using equations 2 and 3.

The three techniques were implemented in ConfidenceLimitsIterative.nb. Example results using the three techniques follow.

<u>Example 1: For n = 20, x = 2, 95% confidence</u>

Using equation 3, $UCL_{0.05}(2, 20)$ is 0.283.

Using the CDF approach, the p for which CDF(BinDist(20, p), 2) equals 0.05 is 0.283.

<u>Example 2: For n = 13, x = 0, 90% confidence</u>

Using equation 3, $UCL_{0.1}(0, 13)$ is 0.162 .

Using the CDF approach, the p for which CDF(BinDist(13, p), 0) equals 0.1 is 0.162.

Using equation 2, $UCL_{0.1}(0, 13)$ is 0.162.

5. CONFIDENCE LIMITS FOR THE HYPERGEOMETRIC DISTRIBUTION

For sampling without replacement, UCL_γ depends on the confidence desired, the size of the sample, the number of defective units in the sample, the size of the population, and the number of defects in the population. Let $UCL_\gamma(x, n, D, N)$ denotes UCL_γ for confidence $(1-\gamma)100\%$ for a sample of size n with x defective units for a population N with D defective units.

For the case of no defective units in the sample, SNL documentation for weapons reliability discusses two techniques.

The first SNL technique utilizes a smoothing process. [Muller] The referenced report states that the smoothing technique "adds one more defective unit"; this appears to cause problems as discussed later. For a specific population N, the referenced report develops a formula for finding the minimum sample size n with no defective units required to provide "90% confidence for 95% reliability". This is called a "90/95" sample size in the report. $\gamma = 0.1$. Using the nomenclature of section III, $\mathbb{R}$ is 0.95, so $1 - UCL_{0.1} = 0.95$. $UCL_{0.1} = 0.05$, so $\mathbb{P} = 0.05$. Since $D = (1 - \mathbb{R})N$, $D= 0.05N$. n is the minimum value that satisfies:

$$\frac{\Gamma(0.95N)\Gamma(N-n+1)}{\Gamma(N+1)\Gamma(0.95N-n)} \leq 0.10 \quad \text{(Eqn. 4)}$$

where Γ is the gamma function. n is found by iteration.

With no defective units in the sample, this technique can be generalized for any confidence and reliability $\mathbb{R}$. D is $(1 - \mathbb{R})N$. For a given $\mathbb{R}$ and γ, n is the minimum value that satisfies

$$\frac{\Gamma(RN)\Gamma(N-n+1)}{\Gamma(N+1)\Gamma(RN-n)} \leq \gamma \quad \text{(Eqn. 5)}$$

n is found by iteration.

The second SNL technique uses the PDF for the hypergeometric distribution. [SNL Weapon Reliability Guide] It also is restricted to no defective units in the sample. To find the minimum n required for given N, γ, and $\mathbb{R}$, the following technique can be used. Note that D is $(1 - \mathbb{R})$ N rounded to the nearest integer. For the hypergeometric distribution HyperDist(n, D, N), find the minimum n for which PDF(HyperDist(n, D, N), 0) does not exceed γ. D/N is $UCL_\gamma(0, n, D, N)$ which is $\mathbb{P}$; $\mathbb{R} = \mathbb{1} - \mathbb{P}$. Iteration is used to find the minimum n. For example, for a population of 100, for 90% confidence, for 95% reliability find the minimum n for which PDF(HyperDist (n, 5, 100), 0) exceeds 0.10.

The second SNL technique for the hypergeometric distribution is analogous to the CDF technique for the binomial distribution previously discussed; for the special case of no defects in the sample (x of 0) the CDF is the PDF.

The second SNL technique can be extended for any number of defective units in the sample by using the CDF for the hypergeometric distribution. This extension is referred to as the third technique. To find the minimum n given a specific x required for given N, γ, and $\mathbb{R}$, the following technique can be used. Note that D is (1 - $\mathbb{R}$) N rounded to the nearest integer. For the hypergeometric distribution HyperDist(n, D, N), find the minimum n for which CDF(HyperDist(n, D, N), x) does not exceed γ. D/N is UCL_{γ}(x, n, D, N) which is $\mathbb{P}$; $\mathbb{R} = \mathbb{1} - \mathbb{P}$. Iteration is used to find the minimum n. For example, for 3 failures in the sample, for a population of 100, for 95% confidence, for 95% reliability find the minimum n for which CDF(HyperDist (n, 5, 100), 3) exceeds 0.05.

The three techniques were implemented in ConfidenceLimitsIterative.nb. Example results using the three techniques follow.

<u>Example 3: N = 100, confidence of 90%, reliability of 95%, no defective items in sample</u>

D is (1 - $\mathbb{R}$) N = 5.

Using the first SNL technique (equation 5), n is 32.

Using the second SNL technique or the third technique, n is 37.

<u>Example 4: N = 20, confidence of 90%, reliability of 90%, no defective items in sample</u>

D is (1 - $\mathbb{R}$) N = 2.

Using the first SNL technique (equation 5), n is 11.

Using the second SNL technique or the third technique, n is 14.

As indicated in these two examples, the two SNL techniques do not provide the same result. The difference between the two techniques is due to the first technique implicitly "adding an additional defective unit" as discussed on the reference for that technique. [Muller] This can be verified as follows. Consider example 3. For N of 100 and 95% reliability, D is 5. If we "add an additional defective unit" D is 6, and the reliability decreases to 94%. With a reliability of 94%, the second and third techniques indicate a sample size of 32 is required, the sample size from the first SNL technique.

Consider example 4. For N of 20 and 90% reliability, D is 2. If we "add an additional defective unit" D is 3, and the reliability decreases to 85%. With a reliability of 85%, the second and third techniques indicate a sample size of 11 is required, the sample size from the first SNL technique.

To check the validity of the third technique- the CDF extension of the second SNL technique- to cases with defective items in the sample, results were compared to published graphs of confidence limits for the hypergeometric distribution. [Hypergeometric Graphs] The graphs provide two-sided confidence limits (upper and lower); the upper one-sided confidence limit can also be obtained from the graphs as stated in the reference. For example, 90%, 95%, 99% upper two-sided limits are equivalent to 95%, 97.5%, 99.5% upper one-sided confidence limits, respectively.

Examples comparing results of the third technique to the published graphical results follow.

<u>Example 5: N = 500, D = 110, n = 50, x = 6</u>

> The graph from the reference indicates 95% confidence (90% upper two-sided on the graph) .
>
> The third technique indicates that for N = 500, D = 110, n = 50, x = 6, γ is 0.0468, so the confidence, (1-γ)100% , is 95.3% in close agreement with the result from the graph.

<u>Example 6: N = 10,000, D = 35, n = 2000, x = 1</u>

> The graph from the reference indicates 99.5% confidence (99% upper two-sided on the graph) .
>
> The third technique indicates that for N = 10,000, D = 35, n = 2000, x = 1, γ is 0.00391145, so the confidence is 99.6% in close agreement with the result from the graph.

<u>Example 7: N = 500, D = 20, n = 250, x = 4</u>

> The graph from the reference indicates 99.5% confidence (99% upper two-sided on the graph) .
>
> The third technique indicates that for N = 500, D = 20, n = 250, x = 4, γ is 0.00512136, so the confidence is 99.5% in agreement with the result from the graph.

6. SAMPLE SIZES FOR NUCLEAR WEAPONS RELIABILITY

SNL uses confidence limits for reliability to indicate the required number of system level tests to be performed for nuclear weapons. [Muller] For sampling from the population without replacement and assuming no defects found in the sample, the number of weapons to be tested to meet a required confidence/reliability can be specified. Here, two confidence/reliability requirements are considered: 90/90 and 90/95. 90/90 means 90% confidence that the reliability is 90%; 90/95 means 90% confidence that the reliability is 95%. Using the techniques previously described for confidence for the hypergeometric distribution, Tables 1 and 2 summarize required sample sizes for various population sizes. The results were calculated using the ConfidenceLimitsIterative.nb Mathematica code.

Required sample sizes predicted by both the first and second SNL techniques are provided. As discussed previously, the first SNL technique under-predicts the required sample size.

Table 1.
Required Sample Size for Various Population Sizes
for 90% confidence for 90% Reliability
With No failures in the Sample
Sampling without Replacement

Population Size	Required Sample Size from SNL Second Technique	Required Sample Size from SNL First Technique (INCORRECT)
10	9	7
20	14	11
30	16	13
40	17	14
50	18	16
70	19	17
90	20	18
120	20	19
150	21	20
250	21	21
275	22[a]	21
532 or greater	22	22[a]

[a]The sample size for sampling with replacement required by the binomial distribution is 22.

Table 2.
Required Sample Size for Various Population Sizes
for 90% confidence for 95% Reliability
With No failures in the Sample
Sampling without Replacement

Population Size	Required Sample Size from SNL Second Technique	Required Sample Size from SNL First Technique (INCORRECT)
10	9	8
20	18	14
40	27	21
100	37	32
200	41	37
300	42	40
400	43	41
500	43	42
800	44	43
1000	44	44
1200	45[b]	44
2131 or greater	45	45[b]

[b]The sample size for sampling with replacement required by the binomial distribution is 45.

7. CONCLUSIONS AND RECOMMENDATIONS

This report provides a number of techniques for evaluating upper one-sided confidence limits for reliability based on sampling with and without replacement, using the binomial and hypergeometric distributions, respectively.

For the binomial distribution the SNL technique used for the special case of no failures in the sample agrees with a general technique which apples for any number of failures in the sample. This report also provides a technique applicable to any number of failures in the sample using the CDF of the binomial distribution.

For the hypergeometric distribution, two SNL techniques developed for no defective items in the sample do not agree. The first technique under-predicts the required sample size. This is due to an additional defective item being implicitly added in the first technique.

This report also provides a technique applicable to any number of defective items in the sample using the CDF of the hypergeometric distribution. This technique provides results in agreement with published graphical results of confidence limits for the hypergeometric distribution.

To meet confidence/reliability requirements for nuclear weapons, the sample size for sampling without replacement, assuming no defects found in the sample, for various population sizes is provided.

All the techniques discussed in the report are implemented in the Mathematica code, Confidence Limits Iterative.nb written by the author.

REFERENCES

[Beyer] Beyer, W., H. (ed.), CRC Standard Probabilities and Statistical tables and Formulae, CRC Press, 1991.

[Dougherty] Dougherty, E.R., Probability and Statistics for the Engineering, Computing, and Physical Sciences, 1990, Prentice-Hall.

[Hahn] "Surveillance Sampling History (1949 – 2002)" (OUO), Hahn, R., SNL, April 16, 2009.

[Hypergeometric Graphs] Chung, J. H., and DeLury, D. B., Confidence Limits for the Hypergeometric Distribution, University of Toronto Press, 1950.

[Loescher] E-mail from D. Loescher, SNL, to J. Darby, December 3, 2009.

[Mathematica] Mathematica software, 7.0, Wolfram Research Inc.

[Martz] Martz, H. F. and Waller, R. A., Bayesian Reliability Analysis, 1982, Wiley.

[Muller] Muller, F. W., "Topics in Stockpile Evaluation", SAND93-2199, September 1993.

[SNL Weapon Reliability Guide] "Nuclear Weapon Reliability Evaluation Methodology Guide", First Edition, Sandia National Laboratories.

[Stockpile Surveillance] "Surveillance: History and Implications for a Smaller Stockpile" (OUO), Darby, J. and Pedersen, R., SAND2009-7788P, Oct. 29, 2009.

(THIS PAGE INTENTIONALLY LEFT BLANK)

APPENDIX A. MATHEMATICA CODE

This appendix provides the functions and algorithms programmed in the Mathematica program ConfidenceLimitsIterative.nb which implements all six techniques discussed in the main report. Also, solutions for some of the examples from Sections 4 and 5 of the main report are provided.

Equation 2 for the binomial distribution is represented by the following function defined in Mathematica:

```
UCLNoFailures[ n_, γ_] = 1 - γ^(1/n)
(* SNL technique *)
```

Equation 3 for the binomial distribution is represented by the following function defined in Mathematica:

```
UCLMartz[x_, n_, γ_] =
(1 + (n - x)/((x + 1) Quantile[FRatioDistribution[2 x + 2, 2 n - 2 x], 1 - γ]))^-1
(*Martz 3.19 for one sided *)
```

The third technique for the binomial distribution is implemented using the following function defined in Mathematica:

```
GammaFromBinomialCDF[n_, P_, x_] = CDF[BinomialDistribution[n, P], x]
(* using CDF solve for P such that function value equals γ *)
```

Equation 5 for the hypergeometric distribution is represented by the following function defined in Mathematica:

```
SNLOneIncorrect[R_, Nh_, n_] =
Gamma[R*Nh] * Gamma[Nh - n + 1] / (  Gamma[R*Nh - n] * Gamma[Nh + 1])
(* SNL approach 1, no failures in sample, incorrect model *)
```

The second SNL technique for the hypergeometric distribution is implemented using the following function defined in Mathematica:

```
GammaSNLTwo[n_,Dh_, Nh_] = PDF[HypergeometricDistribution[n, Dh, Nh], 0]
(* SNL approach 2, no failures in sample *)
```

The third SNL technique for the hypergeometric distribution is implemented using the following function defined in Mathematica:

```
GammaFromHyperGeometricCDF[n_, Dh_, Nh_, x_] =
CDF[HypergeometricDistribution[n, Dh, Nh], x]
(* x failures in sample, solve for minimum D such that function
value exceeds γ, probability failure for confidence for γ is D/N *)
```

The algorithms and solutions for selected examples in Sections 4 and 5 of the main report follow.

Example 1 Binomial Distribution: For n = 20, x = 2, 95% confidence

Using equation 3:

```
Clear[x, n, γ, P, R, Nh, Dh, boolTest];
n = 20; (* enter n *)
γ = 0.05;  (* enter γ *)
x = 02; (* enter x *)
Print["Binomial Distribution UCLMartz. With sample size ", n, ", number of failures in sample ",
x, ",
and required confidence ", (1 - γ) 100, "%.  P is ", UCLMartz[x, n, γ],
 ", and R is " , 1 -  UCLMartz[x, n, γ]]
```

Binomial Distribution UCLMartz. With sample size 20 , number of failures in sample 2 , and required confidence 95. %. P is 0.282619 , and R is 0.717381

Using the CDF approach:

```
(* Given n and x, find P for which GammaFromCDF[n, P, x] is <= γ *)
Clear[x, n, γ, P, R, Nh, Dh, boolTest];
n = 20; (* enter n *)
γ = 0.05; (* enter γ *)
x = 2;
Print["Binomial Distribution GammaFromBinomiaUCL. With sample size of ", n, ",
number of failures in sample ", x, ", and required confidence ", (1 - γ) 100, "%.  P is: ",
FindRoot[GammaFromBinomialCDF[n, P, x]  == γ, {P, 0.5}],
 ", and R is: " , 1 -  FindRoot[GammaFromBinomialCDF[n, P, x]  == γ, {P, 0.5}]]
```

Binomial Distribution GammaFromBinomiaUCL. With sample size of 20 , number of failures in sample 2 , and required confidence 95. %. P is: {P→0.282619} , and R is: {1-(P→0.282619)}

Example 2 Binomial Distribution: For n = 13, x = 0, 90% confidence

Using equation 3:

```
Clear[x, n, γ, P, R, Nh, Dh, boolTest];
n = 13; (* enter n *)
γ = 0.1;  (* enter γ *)
x = 0; (* enter x *)
Print["Binomial Distribution UCLMartz. With sample size ", n, ",
number of failures in sample ", x, ", and required confidence ", (1 - γ) 100, "%.
P is ", UCLMartz[x, n, γ],
 ", and R is " , 1 -  UCLMartz[x, n, γ]]
```

Binomial Distribution UCLMartz. With sample size 13 , number of failures in sample 0 , and required confidence 90. %. P is 0.162322 , and R is 0.837678

Using the CDF approach:

```
(* Given n and x, find P for which GammaFromCDF[n, P, x] is <= γ *)
Clear[x, n, γ, P, R, Nh, Dh, boolTest];
n = 13; (* enter n *)
γ = 0.1; (* enter γ *)
x = 0;
Print["Binomial Distribution GammaFromBinomiaUCL. With sample size of ", n, ",
number of failures in sample ", x, ", and required confidence ", (1 - γ) 100, "%.  P is: ",
FindRoot[GammaFromBinomialCDF[n, P, x]  == γ, {P, 0.5}],
 ", and R is: " , 1 -  FindRoot[GammaFromBinomialCDF[n, P, x]  == γ, {P, 0.5}]]
```

Binomial Distribution GammaFromBinomiaUCL. With sample size of 13 , number of failures in sample 0 , and required confidence 90. %. P is: {P→0.162322} , and R is: {1-(P→0.162322)}

Using equation 2:

```
(* For x of 0, given n and γ, Calculate the upper one sided confidence
limit for P parameter for binomial distribution *)
Clear[x, n, γ, P, R, Nh, Dh, boolTest];
n = 13; (* enter n *)
γ = 0.1; (* enter γ *)
Print["Binomial Distribution UCLNoFailures. With sample size ", n, ",
number of failures in sample ", x, " , and required confidence ", (1 - γ) 100, "%.  ", (1 - γ) 100,
"%.
P is: ", UCLNoFailures[ n, γ],
 ", and R is: " , 1 -  UCLNoFailures[ n, γ]]
```

Binomial Distribution UCLNoFailures. With sample size 13 , number of failures in sample x , and required confidence 90. %. 90. %. P is: 0.162322 , and R is: 0.837678

<u>Example 3 Hypergeometric Distribution: N = 100, confidence of 90%, reliability of 95%, no defective items in sample</u>

D is (1 - ℝ) N = 5.

Using equation 5:

```
(* Assumes x = 0. Given Nh, R, and γ, Find required n.  Note: Dh = (1 - R)Nh *)
Clear[x, n, γ, P, R, Nh, Dh, boolTest];
Nh = 100; (* enter required Nh *)
R = 0.95; (* enter required R *)
```

```
γ = 0.1; (* enter γ *)
(* Iterate upward on n starting from 0 until SNLOneIncorrect[R, Nh, n] is less than γ *)
n = 0;
boolTest = False;
While[! boolTest && n <= Nh,
        {If[SNLOneIncorrect[R, Nh, n] <= γ,
        {boolTest = True; Print["INCORRECT  Hypergeometric Distribution SNLOneIncorrect
is ",
        N[SNLOneIncorrect[R, Nh, n]], ".  For population ", Nh, ", with required reliability ", R ,
"%,
        required confidence ", (1 - γ) 100, "%, and No failures in sample.  Minimum sample size
is ", n]},
  { Print["INCORRECT  For n of ", n, " SNLOneIncorrect[R, Nh, n] is ",
  N[SNLOneIncorrect[R, Nh, n]]], n = n + 1}]
 }
]
```

```
INCORRECT  For n of  0  SNLOneIncorrect[R, Nh, n] is  1.
INCORRECT  For n of  1  SNLOneIncorrect[R, Nh, n] is  0.94
INCORRECT  For n of  2  SNLOneIncorrect[R, Nh, n] is  0.88303
INCORRECT  For n of  3  SNLOneIncorrect[R, Nh, n] is  0.828967
INCORRECT  For n of  4  SNLOneIncorrect[R, Nh, n] is  0.777691
INCORRECT  For n of  5  SNLOneIncorrect[R, Nh, n] is  0.729085
INCORRECT  For n of  6  SNLOneIncorrect[R, Nh, n] is  0.683038
INCORRECT  For n of  7  SNLOneIncorrect[R, Nh, n] is  0.63944
INCORRECT  For n of  8  SNLOneIncorrect[R, Nh, n] is  0.598185
INCORRECT  For n of  9  SNLOneIncorrect[R, Nh, n] is  0.559173
INCORRECT  For n of  10  SNLOneIncorrect[R, Nh, n] is  0.522305
INCORRECT  For n of  11  SNLOneIncorrect[R, Nh, n] is  0.487484
…
INCORRECT  For n of  28  SNLOneIncorrect[R, Nh, n] is  0.131067
INCORRECT  For n of  29  SNLOneIncorrect[R, Nh, n] is  0.120145
INCORRECT  For n of  30  SNLOneIncorrect[R, Nh, n] is  0.109992
INCORRECT  For n of  31  SNLOneIncorrect[R, Nh, n] is  0.100564
```

INCORRECT Hypergeometric Distribution SNLOneIncorrect is 0.0918192 . For population 100 , with required reliability 0.95 %, required confidence 90. %, and No failures in sample. Minimum sample size is 32

Using the second SNL approach for the hypergeometric distribution:

```
(* Assumes x = 0. Given Nh, R, and γ, Find required n such that
GammaSNLTwo[n_,Dh_,Nh_] is less than γ,  Note: Dh = (1 - R)Nh *)
Clear[x, n, γ, P, R, Nh, Dh, boolTest];
Nh = 100; (* enter required Nh *)
γ = 0.1; (* enter γ *)
```

```
R = 0.95;(* enter required R *)
Dh = Round[(1 - R) Nh]; (* Dh = (1 - R)Nh  rounded to closest integer*)
(* Iterate upward on n starting at 0 until GammaSNLTwo[n,Dh,Nh] is less than γ *)
n = 0;
boolTest = False;
While[! boolTest && n <= Nh,
        {If[GammaSNLTwo[n, Dh, Nh] <= γ,
        {boolTest = True; Print["Hypergeometric Distribution GammaSNLTwo is ",
        N[GammaSNLTwo[n, Dh, Nh]], " .  For population ", Nh, ", with required confidence ",
(1 - γ) 100,
    "%, and with required reliability ", R 100 , "% (maximum number defects, rounded, is ", Dh,
").
    With No failures in sample required sample size is ", n]},
  { Print["For n of ", n, " GammaSNLTwo[n,Dh,Nh] is ", N[GammaSNLTwo[n, Dh, Nh]]], n =
n + 1}]
  }
 ]
```

```
For n of  0  GammaSNLTwo[n,Dh,Nh] is  1.
For n of  1  GammaSNLTwo[n,Dh,Nh] is  0.95
For n of  2  GammaSNLTwo[n,Dh,Nh] is  0.90202
For n of  3  GammaSNLTwo[n,Dh,Nh] is  0.855999
For n of  4  GammaSNLTwo[n,Dh,Nh] is  0.811875
For n of  5  GammaSNLTwo[n,Dh,Nh] is  0.76959
For n of  6  GammaSNLTwo[n,Dh,Nh] is  0.729085
For n of  7  GammaSNLTwo[n,Dh,Nh] is  0.690304
For n of  8  GammaSNLTwo[n,Dh,Nh] is  0.653191
For n of  9  GammaSNLTwo[n,Dh,Nh] is  0.617691
For n of  10  GammaSNLTwo[n,Dh,Nh] is  0.583752
…
For n of  33  GammaSNLTwo[n,Dh,Nh] is  0.128277
For n of  34  GammaSNLTwo[n,Dh,Nh] is  0.118704
For n of  35  GammaSNLTwo[n,Dh,Nh] is  0.109711
For n of  36  GammaSNLTwo[n,Dh,Nh] is  0.101272
```

Hypergeometric Distribution GammaSNLTwo is 0.0933601 . For population 100 , with required confidence 90. %, and with required reliability 95. % (maximum number defects, rounded, is 5). With No failures in sample required sample size is 37

Using the third approach for the hypergeometric distribution:

```
(* Given Nh, R, x and γ, Find required n such that GammaFromCDF[n, Dh, Nh, x] = γ,
Note: Dh = (1 - R)Nh *)
Clear[x, n, γ, P, R, Nh, Dh, boolTest]
Nh = 100; (* enter required Nh *)
x = 0; (* enter required x must be <= Nh *)
```

```
γ = 0.1; (* enter γ *)
R = 0.95; (* enter required R *)
Dh = Round[(1 - R) Nh]; (* Dh = (1 - R)Nh  rounded to closest integer*)
(* Iterate upward on n starting at 0 until GammaFromHyperGeometricCDF[n, Dh_ Nh, x] is less
than γ *)
n = 0;
boolTest = False;
While[! boolTest && n <= Nh,
          {If[ GammaFromHyperGeometricCDF[n, Dh, Nh, x]  <= γ,
          {boolTest = True; Print["Hypergeometric Distribution GammaFromHyperGeometriCDF
is ",
          N[ GammaFromHyperGeometricCDF[n, Dh, Nh, x]  ], ".  For population ", Nh, ", with
required confidence ", (1 - γ) 100,
      "%, number failures in sample ", x, ", with required reliability ",
      R  100, "% (number defects in population, rounded, is ", Dh, ").  Required sample size is ",
n]},
   { Print["For n of ", n, " GammaFromHyperGeometricCDF[n, Dh, Nh, x] is ",
   N[GammaFromHyperGeometricCDF[n, Dh, Nh, x]] ], n = n + 1}]
  }
 ]
If[! boolTest, Print["#!#!#!# No n exists; x is too large #!#!#!#"]]
```

For n of 0 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 1.
For n of 1 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.95
For n of 2 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.90202
For n of 3 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.855999
For n of 4 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.811875
For n of 5 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.76959
For n of 6 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.729085
For n of 7 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.690304
For n of 8 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.653191
For n of 9 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.617691
For n of 10 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.583752
…
For n of 33 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.128277
For n of 34 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.118704
For n of 35 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.109711
For n of 36 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.101272
Hypergeometric Distribution GammaFromHyperGeometriCDF is 0.0933601 .

For population 100 , with required confidence 90. %, number failures in sample 0 ,
with required reliability 95. % (number defects in population, rounded, is 5).
Required sample size is 37

Example 5 Hypergeometric Distribution: N = 500, D = 110, n = 50, x = 6

Using the third approach for the hypergeometric distribution:

```
(* Given Nh, n, x and γ, Find R such that GammaFromCDF[n, Dh, Nh, x] <= γ,  Note: Dh = (1
- R)Nh  *)
Clear[x, n, γ, P, R, Nh, Dh, boolTest]
Nh = 500; (* enter required Nh *)
n = 50; (* enter required n *)
x = 6; (* enter required x must be <= n *)
γ = 0.05;
(* iterate upward on Dh starting at 1 until GammaFromHyperGeometricCDF[n, Dh, Nh, x] is
less than γ *)
Dh = 1; (* R = N[1 - Dh/Nh] *)
boolTest = False;
While[! boolTest && Dh <= Nh,
        {If[GammaFromHyperGeometricCDF[n, Dh, Nh, x] <= γ,
        {boolTest = True; Print["Hypergeometric Distribution GammaFromCDF is ",
N[GammaFromHyperGeometricCDF[n, Dh, Nh, x] ], ".
        For population ", Nh, ", with required confidence ", (1 - γ) 100,
     "%, number failures in sample ", x, ", and sample size ", n, ".  Reliability is ",
     N[1 - Dh/Nh] 100 , "%, (number defects in population is ", Dh ")"]},
  { Print["For Dh of ", Dh, " GammaFromHyperGeometricCDF[n, Dh, Nh, x] is ",
  N[GammaFromHyperGeometricCDF[n, Dh, Nh, x]]], Dh = Dh + 1}]
 }
]
```

```
For Dh of  1  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 1.
For Dh of  2  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  1.
For Dh of  3  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  1.
For Dh of  4  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  1.
For Dh of  5  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  1.
For Dh of  6  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  1.
For Dh of  7  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  1.
For Dh of  8  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  1.
For Dh of  9  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  0.999998
For Dh of  10  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  0.999994
For Dh of  11  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  0.999984
For Dh of  12  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  0.999964
For Dh of  13  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  0.999928
For Dh of  14  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  0.999868
For Dh of  15  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  0.999771
For Dh of  16  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  0.999624
For Dh of  17  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  0.99941
For Dh of  18  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  0.999108
For Dh of  19  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  0.998696
For Dh of  20  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  0.998148
For Dh of  21  GammaFromHyperGeometricCDF[n, Dh, Nh, x] is  0.997435
```

…
For Dh of 105 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.0660439
For Dh of 106 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.0617367
For Dh of 107 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.0576694
For Dh of 108 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.0538319
For Dh of 109 GammaFromHyperGeometricCDF[n, Dh, Nh, x] is 0.0502143

Hypergeometric Distribution GammaFromCDF is 0.0468071 . For population 500 , with required confidence 95. %, number failures in sample 6 , and sample size 50 . Reliability is 78. %, (number defects in population is 110)

DISTRIBUTION

1	MS0492	Thomas D. Brown	0411 (electronic copy)
1	MS0415	John Darby	0541 (electronic copy)
1	MS0830	Kathleen Diegert	0413 (electronic copy)
1	MS0417	David Fordham	0543 (electronic copy)
1	MS0415	Steven Hatch	0541 (electronic copy)
1	MS0830	Douglas Loescher	0413 (electronic copy)
1	MS0415	Ronald Pedersen	0541 (electronic copy)
1	MS0829	Janet Sjulin	0415 (electronic copy)
1	MS0899	Technical Library	9536 (electronic copy)

Sandia National Laboratories