Designation: E141 – 10

An American National Standard

# Standard Practice for Acceptance of Evidence Based on the Results of Probability Sampling[1]

This standard is issued under the fixed designation E141; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

## 1. Scope

1.1 This practice presents rules for accepting or rejecting evidence based on a sample. Statistical evidence for this practice is in the form of an estimate of a proportion, an average, a total, or other numerical characteristic of a finite population or lot. It is an estimate of the result which would have been obtained by investigating the entire lot or population under the same rules and with the same care as was used for the sample.

1.2 One purpose of this practice is to describe straightforward sample selection and data calculation procedures so that courts, commissions, etc. will be able to verify whether such procedures have been applied. The methods may not give least uncertainty at least cost, they should however furnish a reasonable estimate with calculable uncertainty.

1.3 This practice is primarily intended for one-of-a-kind studies. Repetitive surveys allow estimates of sampling uncertainties to be pooled; the emphasis of this practice is on estimation of sampling uncertainty from the sample itself. The parameter of interest for this practice is effectively a constant. Thus, the principal inference is a simple point estimate to be used as if it were the unknown constant, rather than, for example, a forecast or prediction interval or distribution devised to match a random quantity of interest.

1.4 A system of units is not specified in this standard.

1.5 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

## 2. Referenced Documents

2.1 *ASTM Standards:*[2]

E105 Practice for Probability Sampling of Materials

E122 Practice for Calculating Sample Size to Estimate, With Specified Precision, the Average for a Characteristic of a Lot or Process

E456 Terminology Relating to Quality and Statistics

E1402 Guide for Sampling Design

E2586 Practice for Calculating and Using Basic Statistics

## 3. Terminology

3.1 *Definitions*—Refer to Terminology E456 for definitions of other statistical terms used in this practice.

3.1.1 *audit subsample, n*—a small subsample of a sample selected for review of all sample selection and data collection procedures.

3.1.2 *equal complete coverage result, n*—the numerical characteristic of interest calculated from observations made by drawing randomly from the frame, all of the sampling units covered by the frame.

3.1.2.1 *Discussion*—Locating the units and evaluating them are supposed to be done in exactly the same way and at the same time as was done for the sample. The quantity itself is denoted θ. The equal complete coverage result is never actually calculated. Its purpose is to serve as the objectively defined concrete goal of the investigation. The quantity θ may be the population mean, ($\bar{Y}$), total ($Y$), median ($M$), the proportion ($P$), or any other such quantity.

3.1.3 *frame, n*—a list, compiled for sampling purposes, which designates all of the sampling units (items or groups) of a population or universe to be considered in a specific study. **E1402**

[1] This practice is under the jurisdiction of ASTM Committee E11 on Quality and Statistics and is the direct responsibility of Subcommittee E11.10 on Sampling / Statistics.

Current edition approved May 15, 2010. Published August 2010. Originally approved in 1959. Last previous edition approved in 2003 as E141 – 91 (2003)$^{\varepsilon 1}$. DOI: 10.1520/E0141-10.

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

3.1.4 *probability sample, n*—a sample in which the sampling units are selected by a chance process such that a specified probability of selection can be attached to each possible sample that can be selected. **E1402**

3.1.5 *replicate subsamples, n*—a number of disjoint samples, each one separately drawn from the frame in accord with the same probability sampling plan.

3.1.6 *sample, n*—a group of observations or test results, taken from a larger collection of observations or test results, which serves to provide information that may be used as a basis for making a decision concerning the larger collection. **E2586**

3.1.7 *sampling unit, n*—an item, group of items, or segment of material that can be selected as part of a probability sampling plan. **E1402**

## 4. Significance and Use

4.1 This practice is designed to permit users of sample survey data to judge the trustworthiness of results from such surveys. Practice E105 provides a statement of principles for guidance of ASTM technical committees and others in the preparation of a sampling plan for a specific material. Guide E1402 describes the principal types of sampling designs. Practice E122 aids in deciding on the required sample size.

4.2 Section 5 gives extended definitions of the concepts basic to survey sampling and the user should verify that such concepts were indeed used and understood by those who conducted the survey. What was the frame? How large (exactly) was the quantity *N*? How was the parameter θ estimated and its standard error calculated? If replicate subsamples were not used, why not? Adequate answers should be given for all questions. There are many acceptable answers to the last question.

4.3 If the sample design was relatively simple, such as simple random or stratified, then fully valid estimates of sampling variance are easily available. If a more complex design was used then methods such as discussed in Ref (**1**)[3] or in Guide E1402 may be acceptable. Use of replicate subsamples is the most straightforward way to estimate sampling variances when the survey design is complex.

4.4 Once the survey procedures that were used satisfy Section 5, see if any increase in sample size is needed. The calculations for making it objectively are described in Section 6.

4.5 Refer to Section 7 to guide in the interpretation of the uncertainty in the reported value of the parameter estimate, $\hat{\theta}$, that is, the value of its standard error, $se(\hat{\theta})$. The quantity $se(\hat{\theta})$ should be reviewed to verify that the risks it entails are commensurate with the size of the sample.

4.6 When the audit subsample shows that there was reasonable conformity with prescribed procedures and when the known instances of departures from the survey plan can be shown to have no appreciable effect on the estimate, the value of $\hat{\theta}$ is appropriate for use.

[3] The boldface numbers in parentheses refer to a list of references at the end of this standard.

## 5. Concepts and Procedures of Sampling

5.1 *Probability sampling* is a procedure by which one obtains a result from a selected set of sampling units that will agree, within calculable limits of variation, with the equal complete coverage result. Probability sampling plans include instructions for using either (*1*) prepared tables of random numbers, (*2*) computer algorithms to generate pseudo-random numbers, or (*3*) certifiably honest physical devices to select the sample units so that inferences may be drawn from the test results and decisions may be made with risks correctly calculated by probability theory.

5.1.1 Such plans are defined and their relative advantages discussed in Guide E1402 and Refs (**1-3**).

5.2 *Procedures* must be described in written form. Parties interested in collecting data should agree on the importance of knowing θ and its definition including measurement methods. The frame shall be carefully and explicitly constructed. Every sampling unit in the frame (*1*) has a unique serial number, which may be preassigned or determined by some definite rule and (*2*) has an address—a complete and clear instruction (or rules for its formulation) as to where and when to make the observation or evaluation. Address instructions should refer to concrete clerical materials such as directories, dials of clocks or of meters, ledgers, maps, aerial photographs, etc. Duplicates in the frame shall be eliminated. *N* shall be well established. Random numbers (or a certifiably honest physical random device) shall dictate selection of the sample. There shall be no substitution of one sampling unit for another. The method of sample selection shall permit calculation of a standard error of the estimate. The use of replicate subsamples is recommended (see 5.4). An audit subsample should be selected and processed and any departures from prescribed measurement methods and location instructions noted (see 5.5). A report should list $\hat{\theta}$ and its standard error with the degrees of freedom in the $se(\hat{\theta})$.

5.3 *Parameter Definition*—The equal complete coverage result may or may not be acceptable evidence. Whether it is acceptable depends on many considerations such as definitions, method of test, care exercised in the testing, completeness of the frame, and on other points not to be settled by statistical theory since these points belong to the subject matter, and are the same whether one uses sampling or not. Mistakes, whether in testing, counting, or weighing will affect the result of a complete coverage just as such mistakes will affect the sample result. By a more expensive method of measurement or more elaborate sampling frame, it may be possible to define a quantity, θ', as a target parameter or ideal goal of an investigation. Criticism that holds θ to be an inappropriate goal should demonstrate that the numerical difference between θ and θ' is substantial. Measurements may be imprecise but so long as measurement errors are not too biased, a large size of the lot or population, *N*, insures that θ and θ' are essentially equal.

5.4 *Replicate Subsamples*—When appropriate, separate laboratories should each work on separate replicate subsamples and teams of investigators should be assigned to separate replicate subsamples. This approach insures that the calculated standard error will not be a systematic underestimate. Such subsamples were called interpenetrating in Ref (**4**) where many of their basic properties were described. See Ref (**5**) for further theory and applications.

5.4.1 For some types of material, a sample selected with uniform spacing along the frame (systematic sample) has increased precision over a selection made with randomly varying spacings (simple random sample). Examples include sampling mineral ore or grain from a conveyor belt or sampling from a list of households along a street. If the systematic sample is obtained by a single random start the plan is then a probability sampling plan, but it does not permit calculating the standard error as required by this practice. After dividing the sample size by an integer $k$ (such as $k = 4$ or $k = 10$) and using a random start for each of $k$ replicate subsamples, some of the increased precision of systematic sampling (and a standard error on $k - 1$ degrees of freedom) can be achieved.

5.5 An audit subsample of the survey sample should be taken for review of all procedures from use of the random numbers through locating and measurement, to editing, coding, data entry and tabulation. Selection of the audit subsample may be done by putting the $n$ sample observations in order as they are collected, calculating the nearest integer to $\sqrt{n}$ , or some other convenient integer, and taking this number to be the spacing for systematic selection of the audit subsample. As few as 10 observations may be adequate. The review should uncover any gross departures from prescribed practices or any conceptual misunderstandings in the definitions. If the audit subsample is large enough (say 30 observations or more) the regression of audited values on initial observations may be used to calibrate the estimate. This technique is the method of two-phase sampling as discussed in Ref (**1**). Helpful discussion of an audit appears in Ref (**2**).

5.6 The estimate is a quantity calculated on the $n$ sample observations in the same way as the equal complete coverage result θ would have been calculated from the entire set of $N$ possible observations of the population; the symbol $\hat{\theta}$ denotes the estimate. In calculating $\hat{\theta}$, replicate subsample membership is ignored.

5.6.1 An estimate has a sampling distribution induced from the randomness in sample selection. The equal complete coverage result is effectively a constant while any estimate is only the value from one particular sample. Thus, there is a mean value of the sampling distribution and there is also a standard deviation of the sampling distribution.

5.7 The standard error is the quantity computed from the observations as an estimate of the sampling standard deviation of the estimate; $se(\hat{\theta})$ denotes the standard error.

5.7.1 When θ is the population average of the $N$ quantities and a simple random sample of size $n$ was drawn, then the sample average $\bar{y}$ becomes the usual estimate $\hat{\theta}$, where:

$$\hat{\theta} = \bar{y} = \sum_{i=1}^{n} y_i / n \tag{1}$$

The quantities $y_1$ , $y_2$ , ..., $y_n$ denote the observations. The standard error is calculated as:

$$se\left(\hat{\theta}\right) = se(\bar{y}) = \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2 / n(n-1)} \tag{2}$$

There are $n - 1$ degrees of freedom in this standard error.

5.7.1.1 *Example*—When the observations are:

81.6, 78.7, 79.7, 78.3, 80.9, 79.5, 79.8, 80.3, 79.5, 80.7

then $\bar{y} = 79.90$ and $se(\bar{y}) = 0.32$.

5.7.2 *Finite Population Correction (fpc)*—Multiplying $se(\bar{y})$ by $\sqrt{1 - n/N}$ is always correct when the goal of the survey is to estimate the finite population mean ($\theta = \bar{Y}$). If random measurement error exists in the observations, then θ' based on a reference measurement method may be a more appropriate survey goal than θ (see 5.3). If so, then $se(\bar{y})$ would be further adjusted upward by an amount somewhat less than the downward adjustment of the fpc. Both of these adjustments are often numerically so small that these adjustments may be omitted—leaving $se(\bar{y})$ of Eq 2 as a slight overestimate.

5.7.2.1 *Example*—Using the previous data and if $N = 50$, then $se(\bar{y})$ becomes $se(\bar{y}) = 0.28$ after applying the fpc.

5.7.3 *Proportions and Total Counts*—If the quantity of interest is (*a*) a proportion or (*b*) a total and the sample is simple random then the above formulas are still applicable. A proportion is the mean of zeroes and ones, while the total is a constant times the mean.

5.7.3.1 When θ is taken to be the population proportion ($\theta = P$) then

$$\hat{\theta} = p = \sum y_i / n = a/n \tag{3}$$

where $a$ is the number of units in the sample with the attribute, and

$$se(p) = \sqrt{p(1-p)/(n-1)} \tag{4}$$

5.7.3.2 When θ = the population total ($\theta = Y$) then

$$\hat{\theta} = Np \text{ and } se\left(\hat{\theta}\right) = N \cdot se(p) \tag{5}$$

5.7.3.3 *Example*—If a simple random sample of size $n = 200$ has $a = 25$ items with the attribute then the conclusion is $\hat{\theta} = 0.125$ and $se(\hat{\theta}) = 0.023$ on 199 degrees of freedom.

5.7.4 *Standard Error from Replicate Subsampling*—If θ is a parameter other than a mean or if the sample design is complex, then replicate subsamples should be used in the sample design. Denote the $k$ separate estimates as $\theta_i$ , i = 1, 2, ..., $k$ and denote by $\hat{\theta}$ the estimate based on the whole sample. The average of the $\hat{\theta}_i$ will be close to, but in general not equal to $\hat{\theta}$. The standard error of $\hat{\theta}$ is calculated as:

$$se\left(\hat{\theta}\right) = \sqrt{\sum_{i=1}^{k} \left(\hat{\theta}_i - \bar{\theta}\right)^2 / k(k-1)} \tag{6}$$

where $\bar{\theta}$ is the average of the $\hat{\theta}_i$ . The standard error is based on $k - 1$ degrees of freedom.

5.7.4.1 *Example*—The following estimates of the percent of sales of prescription drugs in the scope of an overpricing suit brought by the State of North Carolina were based on 20 replicate subsamples; each followed a stratified cluster sampling design. The separate estimates were: 6.8, 7.1, 8.4, 9.5, 8.6, 4.1, 3.7, 3.2, 3.8, 5.8, 8.8, 5.0, 7.9, 8.8, 8.4, 8.1, 6.0, 6.3, 4.5, 5.8. The value of $\hat{\theta}$ was 6.74 % and $se(\hat{\theta}) = 0.43$ % on 19 degrees of freedom. Notice that $\bar{\theta} = 6.58$ does not equal $\hat{\theta} = 6.74$. This is because $\hat{\theta}$ is a ratio of two overall averages while $\bar{\theta}$ is the average of 20 ratios.

## 6. Adequacy of Sample Size

6.1 *Deciding on Increasing Sample Size*—Choice of sample size should be made carefully using Guide E1402, in accordance with Practice E122, or on a comparable basis. Since procedures for setting sample size are based on judgments of the variability to be encountered, there is a possibility that the standard error as calculated from the data could greatly exceed that anticipated. It may happen that the time period of interest for the population has passed or for some other reasons it is not possible to take more observations, in which case the following discussion should be ignored. Otherwise, a decision may be made to increase the number of replicate subsamples or even to carry out a census of the universe. Such decisions must be made strictly independent of knowledge of $\hat{\theta}$. For example, in adversarial settings one party may feel the size of $\hat{\theta}$ is inappropriate and will seek to have it changed. Therefore, experimental protocols along with the standard error should be reviewed prior to announcement of the estimate $\hat{\theta}$. Once all parties are satisfied (methods are sound, standard error adequate) then the estimate can be furnished.

6.2 *Increasing Sample Size by Calculating Costs and Losses*— To assume that $\theta = \hat{\theta}$ is to make a judgment that the cost of decreasing $se(\hat{\theta})$ by increasing sample size is greater than the risks stemming from θ not equal to $\hat{\theta}$. If $n$ is to be increased it is necessary to understand the survey costs as well as the costs of inaccuracies in $\hat{\theta}$. Survey costs are determined through ordinary cost accounting procedures. In judging the seriousness of inaccuracy in $\hat{\theta}$ one needs to imagine losses entailed if θ were one standard error below $\hat{\theta}$ and above $\hat{\theta}$. Calculate these two losses and divide the average by two. This result represents roughly the gain to be expected by quadrupling sample size. If the cost of increasing sample size from $n$ to $4n$ is appreciably less than the above gain there is a basis for increasing sample size.

6.2.1 *Example*—The estimate of percentage "drug-in-suit" sales (see 5.7.4.1) was to be used in determining how much drug companies might have to pay to the state of North Carolina. Thus losses from inaccuracy in $\hat{\theta}$ in this example were relatively clear. The base sales of all prescription drugs in North Carolina was 700 million dollars. About 10 % of "drug-in-suit" sales could be judged as overpricing. An initial sample of only four replicate subsamples was taken and the $se(\hat{\theta})$ was found to be 0.7 %. Thus an overstatement by one standard error would represent a loss to the drug companies of $0.007 \times 700 \times 0.1 = 0.49$ million dollars, while an understatement of the same amount would be the same loss to the state of North Carolina. The average is \$490 000 and half of this is \$245 000. Perhaps, from the court's viewpoint, not all of this is loss since what one party overpays, the other gains. Still the survey would have cost approximately \$50 000 to quadruple in size so it was decided to take the total of 20 subsamples reported on in the Example in 5.7.4.1.

## 7. Reporting Results

7.1 *Basic Technical Report*—The estimate of θ should be reported as "$\hat{\theta}$ with a standard error of $se(\hat{\theta})$ on ν degrees of freedom." This form emphasizes the quantity $\hat{\theta}$ which is to be taken in practice to be the value of θ. It also permits the user to rule out values of θ as improbable under the evidence, by simple calculations based on widely available tables of the Student $t$ distribution.

7.2 *Upper and Lower Confidence Bounds on* θ—Values of θ that can be ruled out because they are (*a*) too large or (*b*) too small, can be calculated as follows:

$$(a)\ \text{Upper bound } \theta(U) = \hat{\theta} + t_{\alpha}(\nu)\, se(\hat{\theta}), \text{ or} \quad (7)$$

$$(b)\ \text{Lower bound } \theta(L) = \hat{\theta} - t_{\alpha}(\nu)\, se(\hat{\theta}) \quad (8)$$

where $t_{\alpha}(\nu)$ is the value from the Student $t$ distribution such that 100α percent of the distribution exceeds $t_{\alpha}(\nu)$. A hypothesized value of θ equal to or larger than θ(*U*) would be rejected by the sample evidence at the α level of significance. For values of $t$ see, for example, Ref (**6**).

7.2.1 *Example*—For the percent drug-in-suit data, a lower bound with 5 % level of significance is found as:

$$\theta(L) = 6.74 - 1.729 \times 0.43 = 6.00, \quad (9)$$

where 1.729 is $t_{0.95}$ from (**6**) with ν = 19.

7.2.2 *Example*—For the percent condition estimate, a 95 % confidence interval would be:

$$79.9 - 2.262 \times 0.32 \text{ to } 79.9 + 2.262 \times 0.32, \text{ or} \quad (10)$$

from 79.2 to 80.6, where 2.262 = $t_{0.975}$ from (**6**) with ν = 9.

7.3 *Three Sigma Limits*—The extreme variation of an estimate (from a probability sample) can often be placed at an interval of three standard deviations above or below the sample result. When the sample is of sufficient size, only 27 out of 10 000 intervals so calculated would not be expected to cover the universe value. Table 1 shows values for $t_{\alpha/2}$ where α = 0.0027, and gives some idea of the effect of having to estimate the standard deviation rather than using previous knowledge of it.

7.3.1 *Example*—For the percent condition estimate, three sigma limits would be set at:

$$79.90 - (4.09 \times 0.32) \text{ to } 79.90 + (4.09 \times 0.32) \text{ or} \quad (11)$$

$$\text{from } 78.59 \text{ to } 81.21$$

7.4 *Nonnormality of the* θ *Distribution*—If any one of the observations is very much smaller or larger than the rest it

**TABLE 1 Student *t* Values Required for Use with a Standard Error on ν Degrees of Freedom to Attain α = 0.0027**

| ν | $t_{\alpha/2}(\nu)$ | ν | $t_{\alpha/2}(\nu)$ | ν | $t_{\alpha/2}(\nu)$ |
|---|---|---|---|---|---|
| 1 | 235.78[A] | 11 | 3.85 | 21 | 3.40 |
| 2 | 19.21 | 12 | 3.76 | 22 | 3.38 |
| 3 | 9.22 | 13 | 3.69 | 23 | 3.36 |
| 4 | 6.62 | 14 | 3.64 | 24 | 3.34 |
| 5 | 5.51 | 15 | 3.59 | 25 | 3.33 |
| 6 | 4.90 | 16 | 3.54 | 26 | 3.32 |
| 7 | 4.53 | 17 | 3.51 | 27 | 3.30 |
| 8 | 4.28 | 18 | 3.48 | 28 | 3.29 |
| 9 | 4.09 | 19 | 3.45 | 29 | 3.28 |
| 10 | 3.96 | 20 | 3.42 | 30 | 3.27 |
| | | | | 40 | 3.20 |
| | | | | 50 | 3.16 |
| | | | | ∞ | 3.00 |

[A] When used to calculate an exact three sigma interval, this value is $t_{\alpha/2}(\nu) = 235.80$ for an exact α/2 = 0.001349898 and ν = 1.

should be investigated. If there is marked asymmetry in the distribution of the observations (for example, there are apparent outliers on one side of the average), be cautious in trusting the realism of α when calculating bounds and confidence limits. If the following estimate of skewness is not larger in absolute value than 0.3, then the change in 100α will likely be less than 1 % due to skewness.

$$g_1 = \frac{n\sum_{i=1}^{n}\left(\hat{\theta}_i - \bar{\theta}\right)^3}{(n-1)(n-2)\left[se\left(\hat{\theta}\right)\right]^3} \tag{12}$$

7.4.1 *Example*—For the 20 replicate subsample estimates of proportion "drug-in-suit," $g_1 = -0.054$, which is far from critical. The value 0.3 in the above rule is a relaxed form of that given on page 42 of Ref (**1**).

7.5 *Bounds on a Proportion in a Large Population When Zero is Observed in a Sample*—It can happen, after observing a random sample of size *n*, that $a = 0$; that is, there are no observations showing the attribute among the *n*. In this case an upper bound with level of significance α is computed as:

$$\theta(U) = 1 - \alpha^{1/n} \tag{13}$$

7.5.1 *Example*—No observations showing the attribute are observed in a sample of size $n = 18$. For $\alpha = 0.05$, $\theta(U) = 0.15$. Any values of the population proportion less than 15 % cannot be ruled out.

7.6 *Bounding Proportions Near Zero in Finite Populations*—For a finite population of size *N*, of which *A* sampling units have the attribute of interest, sampling probabilities are calculated from the hypergeometric distribution, which is the sampling distribution of *a* when the sample is simple random. The full equation is:

$$Pr(a \text{ items with the attribute among } n \text{ in the sample}) \tag{14}$$

$$= \frac{A!(N-A)!(N-n)!\,n!}{a!(A-a)!\,(n-a)!\,(N-A-n+a)!\,N!}$$

7.6.1 An upper bound for the proportion or number in the population with the attribute, given that none are observed in the sample, can be furnished. The first step is to find the probability for $a = 0$ as a sequence of *A* alternating multiplications and divisions where (from Eq 14):

$$Pr(a = 0) = \frac{(N-n)(N-n-1)\ldots(N-n-A+1)}{N(N-1)\ldots(N-A+1)} \tag{15}$$

The upper 5 % bound for *A* is the largest value for which $Pr(a = 0)$ is greater than 0.05.

7.6.1.1 *Example*—We take the example of $n = 20$ with $a = 0$ and suppose $N = 100$. For $A = 14$, $\Pr(a = 0)$ is found to be (80/100)(79/99)…(67/87) = 0.03413, 0.0443 for $A = 13$, 0.0574 for $A = 12$. Thus an upper 5 % bound on *A* is set at $A(U) = 12.5$ when $N = 100$. The upper bound on the finite population proportion becomes 0.125.

7.6.2 Upper and lower bounds for the proportion or number in the population with the attribute, given a small number observed in the sample, can also be derived. Probabilities for $a = 1$, $a = 2$, etc. can be obtained in succession from the probability for $a = 0$.

$$\Pr(a = 1) = \frac{A \cdot n}{(N-A-n+1)\cdot 1}\Pr(a = 0), \tag{16}$$

$$\Pr(a = 2) = \frac{(A-1)(n-1)}{(N-A-n+2)\cdot 2}\Pr(a = 1), \tag{17}$$

$$\Pr(a = 3) = \frac{(A-2)(n-2)}{(N-A-n+3)\cdot 3}\Pr(a = 2), \text{ etc.} \tag{18}$$

7.6.2.1 *Example*—Suppose that there are $N = 800$ items in a lot and let *A* be the unknown number of items with the attribute and that an upper bound with a 97.5 % coefficient ($\alpha = 0.025$) is needed for *A* based on a sample of size $n = 200$. For the case of $A = 31$, $n = 200$ and $N = 800$, Pr(a = 0) = 0.0001096913978. Multiplying by (31/570) and (200/1) brings us to $Pr(a = 1)$; further multiplying by (30/571) and (199/2) gets us to $Pr(a = 2)$; and, finally, multiplying by (29/572) and (198/3) produces $Pr(a = 3) = 0.02087$. Adding these four results gives 0.02841 which is above 0.025. Setting $A = 32$, the chance of observing $a = 3$ or less is 0.02306. The bound itself is set at the half-integer $A(U) = 31.5$ since larger values can be ruled out at the $\alpha = 0.025$ level.

7.6.2.2 *Example*—A lower bound on *A* could also be found for the previous example by trial and error after setting $A = 3$, $A = 4$, and so forth until the probability of 3 or more first exceeds α. When we set $A = 3$ the probability of $a = 3$ becomes 0.0154 so that $A = 3$ can be ruled out by the evidence at the $\alpha = 0.0154$ level of significance. However, when we set $A = 4$ the probability of getting $a = 3$ or $a = 4$ is found to be 0.0503 and so $A(L) = 3.5$ becomes the lower bound.

## 8. Keywords

8.1 audit subsample; confidence limits; estimate; equal complete coverage; finite population correction; hypergeometric distribution; interpenetrating subsamples; population parameter; probability sampling; replicate subsamples; sample size; sampling distribution; sampling frame; sampling unit; skewness; standard error

## REFERENCES


**(1)** Cochran, William G., *Sampling Techniques*, Third Edition, John Wiley and Sons, New York, NY, 1977.

**(2)** Deming, W. Edwards, *Sample Design in Business Research*, John Wiley and Sons, Inc., New York, NY, 1960.

**(3)** Yates, Frank, *Sampling Methods for Censuses and Surveys*, Fourth Edition, MacMillan Publishing Co., Inc., New York, NY, 1981.

**(4)** Mahalanobis, P. C., "Recent Experiments in Statistical Sampling in the Indian Statistical Institute," *Journal of the Royal Statistical Society*, 109:325–370, 1946.

**(5)** Wolter, K., Introduction to Variance Estimation, Second Edition, Springer, New York, NY, 2007.

**(6)** Dixon, W. J., and Massey, F. J, *Introduction to Statistical Analysis*, Fourth Edition, McGraw-Hill, Inc., New York, NY, 1983.