10 Sample Surveys I: Basic Principles & Methods

10.1 Introduction

We consider situations in which the population is concrete and definite and the problem is to obtain some desired information about it. This is done by carrying out an appropriate survey, and performing an appropriate statistical analysis of the data.

The typical examples of the sorts populations that we would like to study and analyze include:

(i) All 18 year olds in a metropolitan area;
(ii) All wheat producers in Northumbria;
(iii) All ‘families’ in the U.K.
(iv) All elderly people who have consulted a doctor over a relevant period.

10.2 Basic Concepts

Target population
The total (finite) population of individuals about which we require information; issues such as content, location, and time may need to be considered.

Study Variables
These are the aspects of the population that we wish to measure, and will define the type(s) of measurement to be taken; usually some aggregate feature of the population, e.g. total amount of debt, average age, proportion of target population that have more than one car.

Sampling Units
Entities that could potentially be included in the sample. May not necessarily be the same as the individuals in the population. For example, could take a sample of addresses in a certain area in order to gain access to information on families that may be living there: thus the addresses are the sampling units, whereas the families are the individuals in the population that we are really interested in.

Sampling Frame
This consists of the set of all sampling units.

Selection Process
In general, we do not examine the entire population, but look at a sample, bearing in mind considerations of accuracy, speed, cost etc. Moreover, we may have a huge population or destructive measurements (e.g. measuring lifetime of batteries, cooking time of a packaged meal etc.): thus selection of survey material needs to be considered carefully.
10.3 Types of sample

10.3.1 Non-probability Samples

Convenience Samples
Criterion is accessibility:
e.g. handing out questionnaires at an underground tube station.

Judgement Samples
Someone exercises deliberate, subjective choice in drawing what they regard as a 'representative' sample: e.g. surveying the members in a 'key' marginal constituency in order to gauge opinion about the government in the population as a whole.

Quota Samples
Combination of judgement and accessibility. Technique is used to combat problems of non-representativity and non-response amongst some key groups of the population. The numbers needed from different groups in order to form a representative sample are determined (possibly through the use of some form of statistical design) giving rise to the quotas for each group. Survey is conducted until the quotas have been filled or 'achieved'.

For e.g., in trying to assess the perceptions people have about a certain sports drink, then we could interview people who go jogging on a regular basis, or take part in long distance races (≥5000km). It may not be easy to access such people (since they’re always on the move). Extra effort may then have to be taken to ensure that at least a certain proportion of the entire sample comes from this group.

10.3.2 Probability Samples

The above methods may lead to unrepresentative samples but the main problem is that we have no way of measuring 'representativeness', or assessing accuracy of our estimates. We need an element of 'randomness' so that we draw samples according to some imposed probability mechanism. In general, we then need to devise a sampling scheme which is economical and easy to operate, yields unbiased estimates, and minimizes the effects of sampling variation.
Possible schemes include:

**Simple Random Sampling**

**Stratified Random Sampling**
In this set up:
- the population is divided into a number of parts called *strata*;
- a sample is drawn independently in each part (or *stratum*);
- variability between strata is eliminated from the sampling error of resulting estimate(s).

**Cluster Sampling**
Employed almost exclusively for administrative convenience, thus reducing sampling costs. Sampling frame consists of *groups* of population members. Groups may be selected at random and all individuals within the selected groups are included in the sample.
For e.g., take a sample of doctors in an area and include all elderly patients who have seen the doctors over a given time period.

**Multi-stage Sampling**
Again, for administrative convenience, sample at several stages.
For e.g., local education authorities, then schools within local education authorities, then classes within a school, then children within a class.

For cluster and multi-stage sampling, the hope is that loss in potential efficiency is outweighed by reduction in sampling costs and effort.
In stratified sampling, the manner in which a population is stratified may be conditioned by administrative factors but the major interest is in using our knowledge about the population being sampled to produce more efficient estimators of population characteristics.

In the rest of our discussion, we focus attention on *probability samples* only.

*Probability samples* are characterized by the following:
- every member of the population has a known probability of being included in the sample;
- the sample is drawn by some method of random selection consistent with these probabilities;
- account is taken of these probabilities of selection in making the estimates from the sample.

In our construction, it is assumed that we have overcome problems of (a) definition of the sampling units, (b) establishing a sampling frame and a random selection process, (c) measurement, questionnaire design, etc.\(^1\)

---

\(^1\)Refer to Barnett for further discussion.
10.4 Example: A Statistics Class

Barnett(1991)

<table>
<thead>
<tr>
<th>ROW</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>15</td>
<td>13</td>
<td>F</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>23</td>
<td>19</td>
<td>M</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>25</td>
<td>16</td>
<td>M</td>
<td>28</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>38</td>
<td>31</td>
<td>M</td>
<td>33</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>25</td>
<td>25</td>
<td>F</td>
<td>23</td>
</tr>
</tbody>
</table>

The table displays data for an actual population of 25 students in a statistics class. The co-ordinates of each cell of the table correspond to where students are seated in the teaching room. Each cell contains information on Gender, Height $Y_i$ (in c.m., in excess of 150 c.m.), and Weight $Z_i$ (in k.g., in excess of 45 k.g.), in that order.

This population is small enough for us to actually compute some of the characteristics of interest. For e.g., the proportion of men, $\pi$, is equal to $15/25=0.6$. The mean height, $\bar{Y}$, is 23.9, and the mean weight, $\bar{Z}$, is 20.9.

Let us try to estimate the value of $\bar{Y}$ using a sample of size 5.

Suppose that ROW 1 were easy to access from the point of view of the lecturer. Then we could view the students in that row as constituting a convenience sample. In such a case

$$\bar{y}_{conv} = (15 + 7 + 13 + 5 + 15)/5 = 11$$

which is low compared to 23.9. But we see that all members of ROW 1 are female.

We could try to remedy this situation by drawing a judgement sample, where we try to ensure that there are reasonable numbers of both gender in the sample. Let us sample the students along the diagonal 1A to 5E. Then

$$\bar{y}_{judge} = (15 + 23 + 20 + 38 + 41)/5 = 27.4$$

Finally, let us construct a simple random sample of size 5.

Labelling the students from 1 to 25 starting from 1A, row by row, then a random selection can be found as follows:

```r
> sample(1:25, size=5)
[1] 3 18 8 20 4
```

Thus, our sample consists of 1C, 4C, 2C, 4E, and 1D, with corresponding sample mean

$$\bar{y} = (13 + 23 + 18 + 30 + 5)/5 = 17.8$$

A different simple random sample of size 5 would give a different value for $\bar{y}$, in general. However, we can say something about the precision of the estimate (which improves as the
sample size $n$ increases). On the other hand, we can say nothing about the precision for $\bar{y}_{\text{conv}}$ and $\bar{y}_{\text{judge}}$: the corresponding sampling procedures would yield the same estimates each time for the population given in the above table. In reality, knowledge of $\bar{Y}$ would be unavailable, and thus there would be no way to gauge the ‘precision’ of the non-probability sample estimates.

10.5 Simple Random Sampling

Suppose that the population has $N$ members and that we are interested in the values taken by some variable $Y$: thus we consider $Y_1, Y_2, \ldots, Y_N$. In general, we are interested in population characteristics defined with reference to the $\{Y_i\}$, which include:

- population total $\sum_{i=1}^{N} Y_i$
- population mean $\frac{1}{N} \sum_{i=1}^{N} Y_i$
- proportion of members of population which fall into some category defined w.r.t. to the $\{Y_i\}$.

We draw a sample of size $n$, where
(a) sampling units are chosen without replacement,
(b) each sampling unit has an equal chance of being chosen by the sampling process.

Thus we need a rule for: (i) drawing the sample (SRS), (ii) making estimates, and we will be interested in questions of bias and precision. Note that even if $N$ is small, the number of possible samples is large and so one might expect considerable sampling variation (e.g. if $N = 40$ and $n = 3$, then the total number of possible unordered samples is $\binom{40}{3} = \frac{40!}{3!37!} = 9880$).

However, we shall consider ordered samples for the purposes of our analysis. Actually, the total number of ordered samples is given by

$$N \times (N-1) \times \ldots \times (N-(n-1))$$

which shall be denoted by $N^{(n)}$.

For each sampling unit in the population, the probability that it will be included in a sample of size $n$ is

$$\frac{n(N-1)(n-1)}{N^{(n)}} = \frac{n}{N}.$$ 

Let $S(n, N)$ be the set of all ordered samples that could be obtained in a single draw of size $n$ from a population of size $N$.

We shall work with the quantities $g = g(y_1, \ldots, y_n)$ and $k = k(y_1, \ldots, y_n)$, which are statistics, constructed from the mappings $g, k: \mathbb{R}^n \mapsto \mathbb{R}$. 

5
Definition 10.1 (Expectation, Variance, Covariance for finite populations)

\[
E[g] = \overline{G} := \frac{1}{N(n)} \sum_{S(n,N)} g(y_1, \ldots, y_n)
\]

\[
\text{var}(g) := E[(g - \overline{G})^2]
\]

\[
\text{cov}(g, k) := E[(g - \overline{G})(k - \overline{K})]
\]

Proposition 10.2 (Properties of the Expectation and Variance Operators)
Suppose that \( C \) is a real constant. Then

\[
E[g + k] = E[g] + E[k]
\]

\[
E[Cg] = CE[g]
\]

\[
\text{var}(g + k) = \text{var}(g) + \text{var}(k) + 2 \text{cov}(g, k)
\]

\[
\text{var}(Cg) = C^2 \text{var}(g)
\]

If the sample size \( n = 1 \), then \( N^{(1)} = N \) and

\[
E[y_1] = \frac{1}{N} (Y_1 + \ldots + Y_N) =: \overline{Y}
\]

\[
\text{var}(y_1) = E[(y_1 - \overline{Y})^2] = \frac{1}{N} \sum_{i=1}^{N} (Y_i - \overline{Y})^2 =: \sigma^2
\]

Now consider drawing a sample of arbitrary size \( n \). Then we can compute the expectation and variance of the \( i \)-th member of the sample as follows:

\[
E[y_i] = \frac{(N-1)^{(n-1)}}{N(n)} \sum_{j=1}^{N} Y_j = \bar{Y}, \quad i = 1, \ldots, n
\]

\[
\text{var}(y_i) = \frac{(N-1)^{(n-1)}}{N(n)} \sum_{j=1}^{N} (Y_j - \overline{Y})^2 = \frac{1}{N} \sum_{j=1}^{N} (Y_j - \overline{Y})^2 = \sigma^2, \quad i = 1, \ldots, n.
\]

We will make repeated use of the following two results.

Proposition 10.3 (A useful result on the expectation of a sum)
Suppose that \( g_0(\cdot) \) is a real-valued function. Then

\[
E[\sum_{i=1}^{n} g_0(y_i)] = \frac{n}{N} \sum_{i=1}^{N} g_0(Y_i)
\]
Proof

\[
E\left[\sum_{i=1}^{n} g_0(y_i)\right] = \frac{1}{N(n)} \sum_{S(n,N)} \sum_{i=1}^{n} g_0(y_i) = \frac{1}{N(n)} \sum_{i=1}^{n} \sum_{S(n,N)} g_0(y_i)
\]

\[
= \frac{1}{N(n)} \sum_{i=1}^{n} \sum_{j=1}^{N} g_0(Y_j) \times (N-1)^{(n-1)} = \frac{(N-1)^{(n-1)}}{N(n)} \sum_{i=1}^{n} \sum_{j=1}^{N} g_0(Y_j)
\]

\[
= \frac{(N-1)^{(n-1)}}{N(n)} \times n \times \sum_{j=1}^{N} g_0(Y_j) = n \frac{(N-1) \times \ldots \times (N-1-(n-2))}{N \times \ldots \times (N-(n-1))} \sum_{i=1}^{N} g_0(Y_i)
\]

\[
= n \frac{(N-1) \times \ldots \times (N-(n-1))}{N \times \ldots \times (N-(n-1))} \sum_{i=1}^{N} g_0(Y_i) = n \frac{N}{N} \sum_{i=1}^{N} g_0(Y_i).
\]

\[\square\]

Proposition 10.4 (Result on the expectation of a double sum)

Suppose that \(g_1(\cdot, \cdot)\) is a real-valued function. Then

\[
E\left[\sum_{i,j=1 \atop i \neq j}^{n} g_1(y_i, y_j)\right] = \frac{n(n-1)}{N(N-1)} \sum_{i,j=1 \atop i \neq j}^{N} g_1(Y_i, Y_j).
\]

Proof

\[
E\left[\sum_{i,j=1 \atop i \neq j}^{n} g_1(y_i, y_j)\right] = \frac{1}{N(n)} \sum_{S(n,N)} \sum_{i,j=1 \atop i \neq j}^{n} g_1(y_i, y_j)
\]

\[
= \frac{1}{N(n)} \sum_{i,j=1 \atop i \neq j}^{n} \sum_{S(n,N)} g_1(y_i, y_j) = \frac{1}{N(n)} \sum_{i,j=1 \atop i \neq j}^{n} \sum_{r,s=1 \atop r \neq s}^{N} g_1(Y_r, Y_s) \times (N-2)^{(n-2)}
\]

\[
= \frac{(N-2)^{(n-2)}}{N(n)} \sum_{i,j=1 \atop i \neq j}^{n} \sum_{r,s=1 \atop r \neq s}^{N} g_1(Y_r, Y_s) = \frac{(N-2)^{(n-2)}}{N(n)} \times n(n-1) \times \sum_{r,s=1 \atop r \neq s}^{N} g_1(Y_r, Y_s)
\]

\[
= n(n-1) \times \frac{(N-2) \times \ldots \times (N-2-(n-3))}{N \times \ldots \times (N-(n-1))} \sum_{r,s=1 \atop r \neq s}^{N} g_1(Y_r, Y_s)
\]

\[
= n(n-1) \times \frac{(N-2) \times \ldots \times (N-(n-1))}{N \times \ldots \times (N-(n-1))} \sum_{r,s=1 \atop r \neq s}^{N} g_1(Y_r, Y_s)
\]

which, after cancellation, yields the desired result.

\[\square\]

Consider

\[
y = \frac{1}{n} \sum_{i=1}^{n} y_i
\]

7
which is the mean obtained from a simple random sample of size \( n \). (See (10) later for the variance of \( \overline{y} \)).

Using Proposition 10.3, it follows that

\[
E[\overline{y}] = \frac{n}{N} \sum_{i=1}^{N} \frac{Y_i}{n} = \frac{1}{N} \sum_{i=1}^{N} Y_i = \overline{Y}.
\] (5)

Another measure of the overall variation in the data of the entire population is given by

\[
S^2 := \frac{1}{N-1} \sum_{i=1}^{N} (Y_i - \overline{Y})^2
\]

(which is not quite the same as \( \sigma^2 \)).

Next we show that

\[
s^2 := \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \overline{y})^2
\]

is an unbiased estimator of \( S^2 \).

**Proposition 10.5 (\( s^2 \) is unbiased for \( S^2 \))**

\[
E[s^2] = S^2
\] (6)

**Proof**

First, observe that

\[
\sum_{i=1}^{n} (y_i - \overline{Y})^2 = \sum_{i=1}^{n} ((y_i - \overline{y}) + (\overline{y} - \overline{Y}))^2 = \sum_{i=1}^{n} (y_i - \overline{y})^2 + n(\overline{y} - \overline{Y})^2.
\]

Hence,

\[
\sum_{i=1}^{n} (y_i - \overline{y})^2 = \sum_{i=1}^{n} (y_i - \overline{Y})^2 - n(\overline{y} - \overline{Y})^2 = A - B
\]

where \( A = \sum_{i=1}^{n} (y_i - \overline{Y})^2 \) and \( B = n(\overline{y} - \overline{Y})^2 \).

By Proposition 10.3,

\[
E[A] = \frac{n}{N} \sum_{i=1}^{N} (Y_i - \overline{Y})^2 = n\sigma^2.
\]

To obtain \( E[B] \), we need to do a bit more work. It is clear that

\[
\left[ \sum_{i=1}^{N} (Y_i - \overline{Y}) \right]^2 = 0
\]
which implies that
\[ \sum_{i=1}^{N} (Y_i - \bar{Y})^2 + \sum_{i,j=1 \atop i \neq j}^{N} (Y_i - \bar{Y})(Y_j - \bar{Y}) = 0 \]

This yields the identity
\[ \sum_{i,j=1 \atop i \neq j}^{N} (Y_i - \bar{Y})(Y_j - \bar{Y}) = -\sum_{i=1}^{N} (Y_i - \bar{Y})^2. \]  
\( (7) \)

Also, since
\[ n(\bar{y} - \bar{Y}) = \sum_{i=1}^{n} (y_i - \bar{Y}) \]
then
\[ n^2(\bar{y} - \bar{Y})^2 = \sum_{i=1}^{n} (y_i - \bar{Y})^2 + \sum_{i,j=1 \atop i \neq j}^{n} (y_i - \bar{Y})(y_j - \bar{Y}). \]

But since the L.H.S. of the above is equal to \( nB \), then
\[ nE[B] = \frac{n}{N} \sum_{i=1}^{N} (Y_i - \bar{Y})^2 + \frac{n(n-1)}{N(N-1)} \sum_{i,j=1 \atop i \neq j}^{N} (Y_i - \bar{Y})(Y_j - \bar{Y}) \]
using Propositions 10.3 and 10.4 to obtain the 1st and the 2nd terms on the R.H.S. respectively. Hence
\[ E[B] = \frac{1}{N} \sum_{i=1}^{N} (Y_i - \bar{Y})^2 - \frac{(n-1)}{N(N-1)} \sum_{i=1}^{N} (Y_i - \bar{Y})^2 \]
\( (8) \)
\[ = \frac{(N-n)}{N(N-1)} \sum_{i=1}^{N} (Y_i - \bar{Y})^2 \]
\( (9) \)
where the 2nd term on the R.H.S. of (8) is obtained from the identity (7).

Finally,
\[ E[A - B] = \frac{n(N-1) - (N-n)}{N(N-1)} \sum_{i=1}^{N} (Y_i - \bar{Y})^2 = \frac{n-1}{N-1} \sum_{i=1}^{N} (Y_i - \bar{Y})^2. \]

Thus,
\[ E[s^2] = E\left[ \frac{1}{n-1}(A - B) \right] = \frac{1}{N-1} \sum_{i=1}^{N} (Y_i - \bar{Y})^2 \]
which is equal to \( S^2 \). \( \square \)
Remarks 10.6 (Variance in terms of the ‘sampling fraction’)
(i) From (9), it follows that

$$E[B] = E[n(\bar{y} - \bar{Y})^2] = \left(1 - \frac{n}{N}\right) S^2$$

Hence

$$\text{var}(\bar{y}) = E[(\bar{y} - \bar{Y})^2] = \left(1 - \frac{n}{N}\right) \frac{S^2}{n} = (1 - f) \frac{S^2}{n}$$

(10)

where \(f = \frac{n}{N}\) is the sampling fraction.
(ii) Note:
- no distributional assumptions have been made;
- \(f\) is less important than \(n\) when \(N\) is large;
- \(\bar{y}\) has minimum variance in the class of linear unbiased estimators (BLUE).

It also turns out that the covariance between any two members of our sample is negative, since

$$\text{cov}(y_i, y_j) = E[(y_i - \bar{Y})(y_j - \bar{Y})] = \frac{(N-2)(n-2)}{N(N-n)} \sum_{k \neq l}^{N} (Y_k - \bar{Y})(Y_l - \bar{Y})$$

$$= -\frac{1}{N(N-1)} \sum_{k=1}^{N} (Y_k - \bar{Y})^2 = -\frac{S^2}{N}$$

(where the penultimate expression follows from the identity (7)).

Remarks 10.7 (Relative importance of \(f\) and \(n\))
Let us consider the result of (10), by tabulating \((1 - f)\frac{S^2}{n}\) for \(S^2 = 1\) and various values of \(n\) and \(N\).

<table>
<thead>
<tr>
<th>(n)</th>
<th>(500)</th>
<th>(1000)</th>
<th>(5000)</th>
<th>(10000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.198</td>
<td>0.199</td>
<td>0.200</td>
<td>0.200</td>
</tr>
<tr>
<td>10</td>
<td>0.098</td>
<td>0.099</td>
<td>0.099</td>
<td>0.100</td>
</tr>
<tr>
<td>50</td>
<td>0.018</td>
<td>0.019</td>
<td>0.020</td>
<td>0.020</td>
</tr>
<tr>
<td>100</td>
<td>0.008</td>
<td>0.009</td>
<td>0.010</td>
<td>0.010</td>
</tr>
<tr>
<td>200</td>
<td>0.003</td>
<td>0.004</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td>500</td>
<td>0.000</td>
<td>0.001</td>
<td>0.002</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Clearly the standard error of the mean, i.e. the square root of its variance, depends mainly on the sample size \(n\) and only to a minor extent on the fraction being sampled. In practical applications, \(f\) is often omitted when \(f = n/N < 10\%\).

\(^2\)Refer to Barnett for further details.
Remarks 10.8 (Sampling with Replacement)
Suppose that we were to sample with replacement. Denoting the corresponding sample mean by \( \bar{y}' \), then

\[
E[\bar{y}'] = \bar{Y} \tag{11}
\]

\[
\text{var}(\bar{y}') = \frac{\sigma^2}{n} \left( 1 - \frac{1}{N} \right) \frac{S^2}{n}. \tag{12}
\]

Comparing (12) with (10), we see that sampling with replacement is bound to be less efficient: it can be shown that this is due to the negative covariance (obtained earlier).

10.6 Estimation of other useful quantities

10.6.1 Population Totals
Consider the population total

\[
T = \sum_{i=1}^{N} Y_i.
\]

We could estimate \( T \) by \( \hat{T} = N \bar{y} \), since

\[
E[\hat{T}] = N \times E[\bar{y}] = N \times \bar{Y} = \sum_{i=1}^{N} Y_i = T \tag{13}
\]

\[
\text{var}(\hat{T}) = N^2 \times (1 - f)S^2 / n = (1 - f)N^2 S^2 / n. \tag{14}
\]

10.6.2 Population Proportions
Here, we want to estimate the proportion of the population that possess a certain attribute. To this end, define \( X_i \) to be equal to 1 if the \( i \)-th individual possesses the attribute, and equal to 0 if the \( i \)-th individual does not possess the attribute.

Let \( R \) be the number of individuals in the population of size \( N \) that possess the attribute, i.e. \( R = \sum_{i=1}^{N} X_i \).

Let \( \pi \) be the proportion of the population that possess this attribute, i.e. \( \pi = R/N \).

Let \( x_i \) take the value 1, if the \( i \)-th member of the s.r. sample (obtained without replacement) of size \( n \) possesses the characteristic, and let it take the value 0 otherwise.

Let \( r \) be the number of members of the sample that possess the attribute, i.e. \( r = \sum_{i=1}^{n} x_i \).

For \( n = 1 \)

\[
E[x_1] = \bar{X} = \pi
\]

\[
\text{var}(x_1) = (1 - f)S^2 = \pi(1 - \pi)
\]
since
\[ S^2 = \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \bar{X})^2 = \frac{1}{N-1} \sum_{i=1}^{N} (X_i^2 - 2\bar{X}X_i + \bar{X}^2) \]
\[ = \frac{1}{N-1} \left\{ \sum_{i=1}^{N} (X_i^2 - 2\bar{X}^2 + N\bar{X}^2) \right\} = \frac{1}{N-1} \left\{ \sum_{i=1}^{N} X_i - N\bar{X}^2 \right\} \]
\[ = \frac{1}{N-1} \left\{ N\bar{X} - N\bar{X}^2 \right\} = \frac{1}{N-1} N\bar{X}(1 - \bar{X}) = \frac{N\pi(1 - \pi)}{N-1} \]

General \( n \)

\[ \bar{x} = \sum_{i=1}^{n} x_i/n \]
\[ s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 = \frac{1}{n-1} n\bar{x}(1 - \bar{x}). \]

Then,
\[ E[\bar{x}] = \bar{X} = R/N = \pi \] \hspace{1cm} (15)
\[ \text{var}(\bar{x}) = (1 - f) \frac{S^2}{n} = \frac{N-n}{N} \times \frac{N\pi(1 - \pi)}{N-1} \times \frac{1}{n} \]
\[ = \frac{(N-n)}{n(N-1)} \pi(1 - \pi). \] \hspace{1cm} (16)