CHRIST UNIVERSITY
BENGALURU, INDIA

**Course Code**: STA405-6

**Course Name**: Sampling Techniques

# CIA – 1 ASSIGNMENT

# Topic - Simple Random Sampling

Submitted by

**Name :** Greeshma S

**Reg No:** 2341436

November – 2025

# A Comprehensive Report on Simple Random Sampling: A Case Study of University Students

## I. Introduction

Statistical inference rests on the foundational principle that we can learn about a large population by studying a smaller, carefully selected subset of it, known as a sample. The method by which this sample is chosen is critical, as a biased selection process can lead to erroneous and misleading conclusions. Among the various probability sampling techniques, Simple Random Sampling (SRS) stands as the most basic and pure form, providing the theoretical underpinning for more complex methods. This report aims to provide a detailed exploration of SRS. Using the hypothetical population of "All undergraduate students at the University of Exemplaria (UE) for the 2023-2024 academic year," we will define the population and sampling frame, delineate the steps for SRS, discuss its advantages and limitations, conceptually differentiate between its two main forms (with and without replacement), and finally, employ the R statistical software to simulate and compare these two methods.

## II. Objectives

The primary objectives of this report are:

1. To clearly define a study population and its corresponding sampling frame.
2. To describe a meticulous, step-by-step procedure for selecting a sample via Simple Random Sampling.

3. To critically evaluate three key advantages and three limitations of SRS in the context of the chosen population.
4. To explain the conceptual difference between Simple Random Sampling With Replacement (SRSWR) and Without Replacement (SRSWOR), with a focus on its implications for variance.
5. To generate a hypothetical population dataset in R, apply both SRSWR and SRSWOR, and compare their results in terms of sample means and variances.

# III. Population and Sampling Frame

### 3.1 Population Definition

The population of interest for this study is unambiguously defined as: **"All individuals who are officially enrolled as full-time or part-time undergraduate students in any degree-granting program at the University of Exemplaria (UE) for the Fall 2023 semester."**

This definition is precise and leaves little room for ambiguity. It includes:

- **Inclusion Criteria:** All undergraduate students, regardless of their year of study (Freshman, Sophomore, Junior, Senior), major, age, nationality, or mode of attendance (full-time or part-time). It covers students across all campuses and faculties of the university.
- **Exclusion Criteria:** It explicitly excludes graduate students (Master's and Doctoral candidates), non-degree seeking students, auditing students, faculty members, administrative staff, and students who have officially withdrawn, are on a leave of absence, or have been admitted for the Spring 2024 semester but are not yet enrolled in Fall 2023.

A clear population definition is paramount as it establishes the universe to which the findings of the study will be generalized. Any inference made from the sample is valid only for this specifically defined group.

### 3.2 Sampling Frame Description

The sampling frame is the actual list, map, or operational device from which the sample is selected. It is the physical representation of the population. For our study, the ideal sampling frame would be the **"Official University of Exemplaria Student Registry Database, snapshot taken as of the end of the add/drop period (Census Date) for the Fall 2023 semester."**

This database is maintained by the University Registrar's office and is used for official academic and administrative purposes. It would typically contain the following fields for each student:

- A unique student identification number (e.g., Student ID)
- Full Name
- Date of Birth
- Academic Program/Major
- Year of Study (e.g., 1, 2, 3, 4)
- Enrollment Status (Full-time/Part-time)
- Email Address
- Campus

**Evaluation and Potential Imperfections of the Sampling Frame:**
While this database is the best available representation of our population, it is crucial to acknowledge that no sampling frame is perfect. Potential issues include:

1. **Coverage Error:** The database might have over-coverage or under-coverage.
    - *Over-coverage* could occur if the list includes students who have unofficially dropped out (stopped attending classes but not formally

withdrawn) or recently graduated in the summer. These are not part of our defined population.

- *Under-coverage* might exist if there are delays in processing the enrollment of newly admitted students, thus excluding eligible population members.

2. **Inaccuracy:** Some data fields might be outdated or incorrect (e.g., a student has changed their major but it is not yet updated in the system).
3. **Duplication:** A student might be listed more than once due to data entry errors, though this is unlikely in a well-maintained university database with a unique student ID.

Before proceeding with sampling, the researcher must work with the Registrar's office to "clean" the frame to the best of their ability, verifying its accuracy and completeness to minimize these potential biases. For the purpose of this report, we will assume the frame has been cleaned and is a near-perfect representation of our target population, consisting of N = 15,000 undergraduate students.

# IV. Steps for Selecting a Sample Using Simple Random Sampling

Selecting a sample via SRS is a systematic process designed to ensure that every possible sample of a given size has an equal probability of being selected. The following is a detailed, step-by-step procedure for our study.

**Step 1: Clearly Define the Target Population and Research Objectives.**
Before any sampling occurs, we must have absolute clarity on *who* we are studying and *why*. As established in Section 3.1, our population is all UE undergraduates. Let's assume our research objective is to "estimate the average annual textbook

expenditure of UE undergraduate students." This objective directly informs what data we need to collect from our sample.

**Step 2: Identify and Obtain the Sampling Frame.**

We procure the cleaned Official Student Registry Database from the University Registrar. This list is our operational sampling frame. We confirm that it contains N = 15,000 unique undergraduate students, each with a unique Student ID.

**Step 3: Determine the Desired Sample Size (n).**

The sample size is a critical decision, balancing precision, confidence, and cost. It is determined through a power analysis or sample size calculation formula. For estimating a population mean, the formula is:

n = (Z^2 * σ^2) / E^2

Where:

- Z is the Z-score corresponding to the desired confidence level (e.g., 1.96 for 95% confidence).
- σ is the estimated population standard deviation (perhaps from a pilot study or previous year's data).
- E is the desired margin of error.

Suppose a pilot study estimated the standard deviation of textbook expenditure (σ) to be $120. For a 95% confidence level and a margin of error (E) of $10, the calculation would be:

n = ( (1.96)^2 * (120)^2 ) / (10)^2 ≈ 553

We may need to apply a finite population correction factor since our population is not infinite, which might adjust the required n slightly downward. For this example, we will target a final sample size of **n = 560 students**.

**Step 4: Assign a Unique Identifier to Every Element in the Frame.**

Each of the 15,000 students already has a unique Student ID. To facilitate the physical process of selection, we will create a numbered list from 1 to 15,000, where each number corresponds to a single, specific student in the database. This creates a direct mapping.

**Step 5: Select the Sample Using a Random Mechanism.**
This is the core of SRS. We must use a method that gives every set of 560 students an equal chance of being chosen. We will use a computer-based random number generator, which is the modern and most efficient standard.

1. Using statistical software like R, we will generate 560 distinct random integers between 1 and 15,000 (for SRS without replacement).
2. The software's random number generator uses algorithms to produce numbers that are statistically indistinguishable from true randomness.
3. The 560 randomly generated numbers are our selected sample. For example, the R command $\text{sample}(1\text{:}15000, 560, \text{replace=FALSE})$ would accomplish this.

**Step 6: Locate and Contact the Selected Elements.**
Using the mapping from Step 4, we identify the 560 students corresponding to the 560 random numbers. We then use their contact information from the database (e.g., university email) to invite them to participate in our survey on textbook expenditures.

**Step 7: Manage Non-Response.**
It is highly unlikely that all 560 selected students will respond. Non-response is a major source of potential bias. Strategies to manage this include:

- Sending multiple follow-up reminders.
- Offering small incentives for participation.
- Conducting a non-response analysis by comparing a small subset of non-respondents to respondents on available characteristics (e.g., year of study) to check for systematic differences.

  The final analysis will be conducted on the actual respondents. The initial sample size was set at 560 anticipating a certain non-response rate to still achieve a sufficient number of completed surveys for analysis.

# V. Advantages and Limitations of SRS in this Study

### 5.1 Advantages

1. **Minimization of Selection Bias:** The most significant advantage of SRS is that the random selection process eliminates conscious or unconscious bias on the part of the researcher. The researcher has no discretion in who is chosen; the computer decides. This prevents the sample from being systematically skewed towards, for example, only students who are active in campus life or only those from a particular faculty, thereby ensuring the sample is representative of the entire population in an unbiased manner.
2. **Ease of Understanding and Implementation:** The concept of SRS is straightforward—a lottery system where every student has an equal chance of being selected. This simplicity makes it easy to explain to university administrators, stakeholders, and the students themselves, which can help in gaining institutional approval and participant trust. Furthermore, with modern software like R, the actual mechanics of selection are simple and quick to execute once the sampling frame is prepared.
3. **Theoretical Foundation for Statistical Inference:** SRS provides the simplest framework for applying statistical theories. Formulas for calculating estimates (like the sample mean), standard errors, and confidence intervals are well-established and uncomplicated when the data comes from an SRS. This allows us to make valid inferences about the population average textbook expenditure from our sample average, with quantifiable measures of uncertainty (e.g., "we are 95% confident that the true average expenditure is between $X and $Y").

### 5.2 Limitations

1. **Requires a Complete and Accurate Sampling Frame:** As discussed in Section 3.2, the quality of the SRS is entirely dependent on the quality of the sampling frame. If the Official Student Registry is missing students (under-coverage) or includes ineligible individuals (over-coverage), the sample will be flawed, and the results will be biased, regardless of the perfect randomness of the selection. Obtaining and cleaning such a frame can be a significant practical hurdle.
2. **Potential for High Cost and Low Efficiency, Especially with a Large, Geographically Dispersed Population:** While UE students are all part of one university, they may be spread across multiple campuses or spend little time on campus. Contacting a simple random sample of 560 students, who are scattered across all majors and years, can be logistically challenging and expensive. An email survey mitigates this, but if in-person interviews were required, the travel costs would be prohibitive. Cluster sampling, where entire classes or dorms are selected first, might be more efficient in such a scenario.
3. **Risk of Under-Representing Small Subgroups:** SRS does not guarantee that small but important subgroups within the population will be proportionally represented in the sample. For instance, if only 2% of undergraduates (300 students) are in a specific major like "Classical Studies," an SRS of 560 might, by chance, select only a handful or even zero students from this group. Any analysis focused on this subgroup would then be impossible or unreliable. If analyzing such subgroups is a key objective, stratified random sampling (where the population is divided into strata like majors, and an SRS is taken from each) would be a more appropriate method.

# VI. Conceptual Difference Between SRSWR and SRSWOR

The distinction between Simple Random Sampling With Replacement (SRSWR) and Without Replacement (SRSWOR) is fundamental and has direct consequences for the properties of the sample.

**Simple Random Sampling Without Replacement (SRSWOR):**
In SRSWOR, once an element (a student) is selected from the population, it is *removed* and is not available for selection again. Imagine writing all 15,000 student IDs on individual slips of paper, placing them in a giant drum, mixing them thoroughly, and drawing 560 slips one by one *without putting any back*. Each student can appear in the sample at most once. This is the most common method used in practice because it is intuitively appealing—we typically do not want to survey the same person multiple times in a single sample.

**Simple Random Sampling With Replacement (SRSWR):**
In SRSWR, after an element is selected, it is *returned* to the population before the next draw. Using the drum analogy, you would draw a slip, record the ID, and then *put the slip back into the drum*and mix it again before the next draw. This means it is possible for the same student to be selected more than once in the same sample.

**Key Conceptual Differences:**

1. **Probability of Selection:** In SRSWOR, the probabilities change with each draw. The probability of a student being selected on the first draw is $1/15000$. If they are not selected, their probability on the second draw becomes $1/14999$. In SRSWR, the probability remains constant for every single draw: $1/15000$. Every draw is independent of the others.
2. **Number of Possible Samples:** There are many more possible samples in SRSWR than in SRSWOR. SRSWOR is about choosing a *set* of 560 distinct students. SRSWR is about a *sequence*of 560 selections where order matters and repetition is allowed. This fundamental difference in the sample space is why the statistical properties differ.

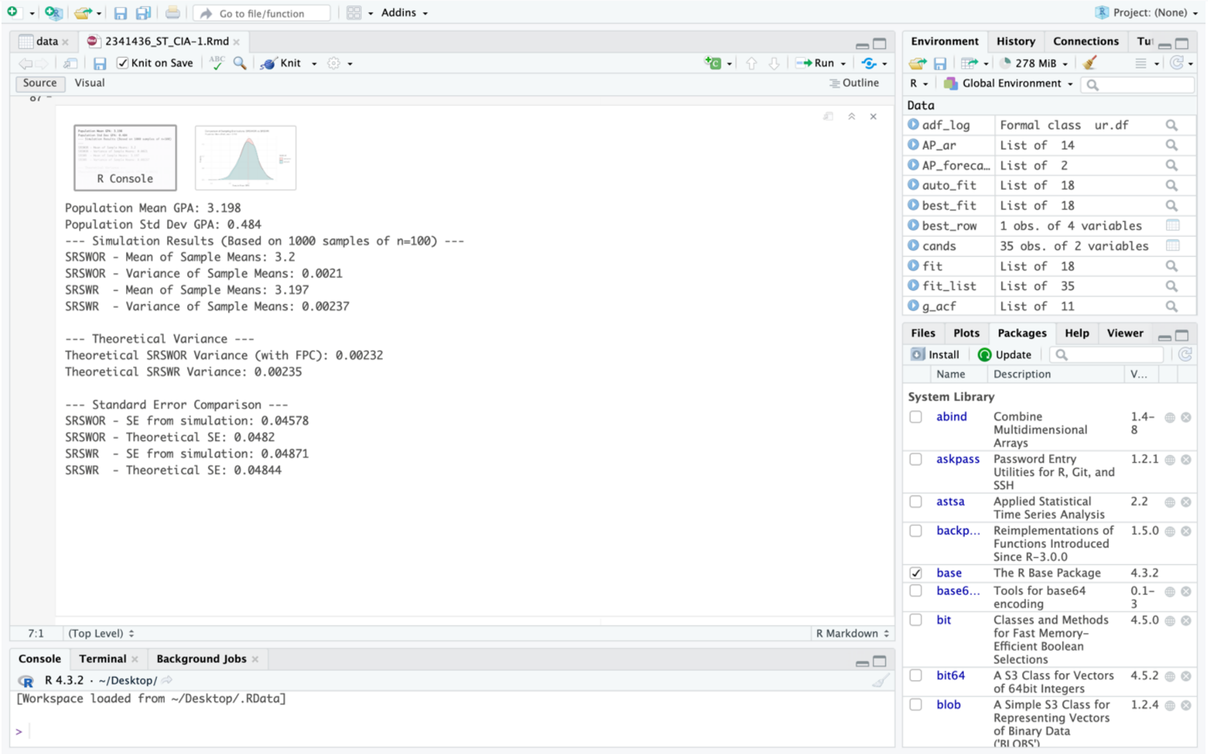
3. **Efficiency and Variance:** This is the most important statistical consequence. **SRSWOR is more efficient than SRSWR.** The sample from SRSWOR contains more "information" about the population because it consists of 560 *distinct* individuals. In SRSWR, if a unit is selected multiple times, the subsequent selections do not provide new information.
   - The variance of an estimator (like the sample mean) is *lower* under SRSWOR.
   - The formula for the variance of the sample mean under SRSWOR includes a **Finite Population Correction (FPC)** factor: $(N - n)/(N - 1)$. Since this factor is less than 1, it reduces the variance.
   - The variance under SRSWR is simply $\sigma^2/n$, which is larger than the SRSWOR variance (which is $(\sigma^2/n) * FPC$).

In summary, SRSWOR is generally preferred in practice as it yields more precise estimates for the same sample size by avoiding the redundancy and loss of information inherent in SRSWR.

# VII. R Simulation: Comparing SRSWR and SRSWOR

We will now use R to simulate a hypothetical population of UE students and draw samples using both methods.

```r
# R Simulation: Comparing SRSWR and SRSWOR

# 7.1 Generating the Hypothetical Population
set.seed(123)
N <- 10000

# Generate hypothetical GPA data (skewed towards higher values, mean ~3.1, sd ~0.5)
population_data <- data.frame(
  StudentID = 1:N,
  GPA = rbeta(N, shape1 = 8, shape2 = 2) * 4 # Scale Beta(8,2) to 0-4 range
)

# Check population parameters
pop_mean <- mean(population_data$GPA)
pop_sd <- sd(population_data$GPA)
cat("Population Mean GPA:", round(pop_mean, 3), "\n")
cat("Population Std Dev GPA:", round(pop_sd, 3), "\n")

# 7.2 Applying SRSWOR and SRSWR
n <- 100
num_sim <- 1000

# Vectors to store results
sample_means_wor <- numeric(num_sim)
sample_means_wr <- numeric(num_sim)

# Simulation loop
for (i in 1:num_sim) {
  # SRS without replacement
  sample_wor <- sample(population_data$GPA, size = n, replace = FALSE)
  sample_means_wor[i] <- mean(sample_wor)

  # SRS with replacement
  sample_wr <- sample(population_data$GPA, size = n, replace = TRUE)
  sample_means_wr[i] <- mean(sample_wr)
}

# 7.3 Comparing the Results
# Create a data frame for plotting
results_df <- data.frame(
  Method = rep(c("SRSWOR", "SRSWR"), each = num_sim),
  SampleMean = c(sample_means_wor, sample_means_wr)
)

# Load ggplot2 for visualization
library(ggplot2)

# Plot the distributions
ggplot(results_df, aes(x = SampleMean, fill = Method)) +
  geom_density(alpha = 0.5) +
  geom_vline(xintercept = pop_mean, linetype = "dashed", color = "red") +
  labs(title = "Comparison of Sampling Distributions: SRSWOR vs SRSWR",
       subtitle = paste("Population Mean (Red Line) =", round(pop_mean, 3)),
       x = "Sample Mean GPA",
       y = "Density") +
  theme_minimal()

# Calculate and print the mean and variance of the sample means
cat("--- Simulation Results (Based on 1000 samples of n=100) ---\n")
cat("SRSWOR - Mean of Sample Means:", round(mean(sample_means_wor), 3), "\n")
cat("SRSWOR - Variance of Sample Means:", round(var(sample_means_wor), 5), "\n")
cat("SRSWR  - Mean of Sample Means:", round(mean(sample_means_wr), 3), "\n")
cat("SRSWR  - Variance of Sample Means:", round(var(sample_means_wr), 5), "\n")

# Theoretical variance for comparison
theoretical_var_wor <- (pop_sd^2 / n) * ((N - n) / (N - 1))
theoretical_var_wr <- (pop_sd^2 / n)

cat("\n--- Theoretical Variance ---\n")
cat("Theoretical SRSWOR Variance (with FPC):", round(theoretical_var_wor, 5), "\n")
cat("Theoretical SRSWR Variance:", round(theoretical_var_wr, 5), "\n")

# Additional comparison: Standard Error
cat("\n--- Standard Error Comparison ---\n")
cat("SRSWOR - SE from simulation:", round(sd(sample_means_wor), 5), "\n")
cat("SRSWOR - Theoretical SE:", round(sqrt(theoretical_var_wor), 5), "\n")
cat("SRSWR  - SE from simulation:", round(sd(sample_means_wr), 5), "\n")
cat("SRSWR  - Theoretical SE:", round(sqrt(theoretical_var_wr), 5), "\n")
```



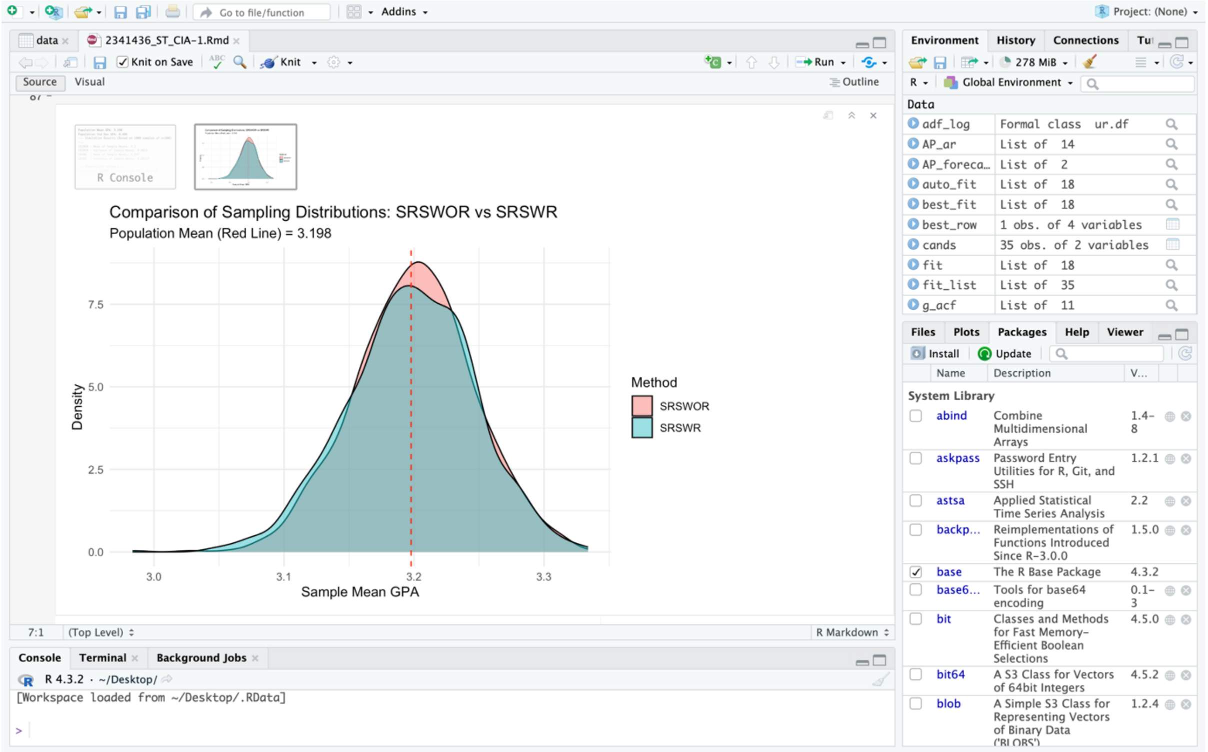
```
Population Mean GPA: 3.198
Population Std Dev GPA: 0.484
--- Simulation Results (Based on 1000 samples of n=100) ---
SRSWOR - Mean of Sample Means: 3.2
SRSWOR - Variance of Sample Means: 0.0021
SRSWR  - Mean of Sample Means: 3.197
SRSWR  - Variance of Sample Means: 0.00237

--- Theoretical Variance ---
Theoretical SRSWOR Variance (with FPC): 0.00232
Theoretical SRSWR Variance: 0.00235

--- Standard Error Comparison ---
SRSWOR - SE from simulation: 0.04578
SRSWOR - Theoretical SE: 0.0482
SRSWR  - SE from simulation: 0.04871
SRSWR  - Theoretical SE: 0.04844
```

```
[Workspace loaded from ~/Desktop/.RData]
>
```

Data

| | |
|---|---|
| adf_log | Formal class ur.df |
| AP_ar | List of 14 |
| AP_foreca… | List of 2 |
| auto_fit | List of 18 |
| best_fit | List of 18 |
| best_row | 1 obs. of 4 variables |
| cands | 35 obs. of 2 variables |
| fit | List of 18 |
| fit_list | List of 35 |
| g_acf | List of 11 |

| | Name | Description | V… |
|---|---|---|---|
| | System Library | | |
| □ | abind | Combine Multidimensional Arrays | 1.4-8 |
| □ | askpass | Password Entry Utilities for R, Git, and SSH | 1.2.1 |
| □ | astsa | Applied Statistical Time Series Analysis | 2.2 |
| □ | backp… | Reimplementations of Functions Introduced Since R-3.0.0 | 1.5.0 |
| ☑ | base | The R Base Package | 4.3.2 |
| □ | base6… | Tools for base64 encoding | 0.1-3 |
| □ | bit | Classes and Methods for Fast Memory-Efficient Boolean Selections | 4.5.0 |
| □ | bit64 | A S3 Class for Vectors of 64bit Integers | 4.5.2 |
| □ | blob | A Simple S3 Class for Representing Vectors of Binary Data ('BLOBS') | 1.2.4 |



Comparison of Sampling Distributions: SRSWOR vs SRSWR

Population Mean (Red Line) = 3.198

```
[Workspace loaded from ~/Desktop/.RData]
>
```

Data

| | |
|---|---|
| adf_log | Formal class ur.df |
| AP_ar | List of 14 |
| AP_foreca… | List of 2 |
| auto_fit | List of 18 |
| best_fit | List of 18 |
| best_row | 1 obs. of 4 variables |
| cands | 35 obs. of 2 variables |
| fit | List of 18 |
| fit_list | List of 35 |
| g_acf | List of 11 |

| | Name | Description | V… |
|---|---|---|---|
| | System Library | | |
| □ | abind | Combine Multidimensional Arrays | 1.4-8 |
| □ | askpass | Password Entry Utilities for R, Git, and SSH | 1.2.1 |
| □ | astsa | Applied Statistical Time Series Analysis | 2.2 |
| □ | backp… | Reimplementations of Functions Introduced Since R-3.0.0 | 1.5.0 |
| ☑ | base | The R Base Package | 4.3.2 |
| □ | base6… | Tools for base64 encoding | 0.1-3 |
| □ | bit | Classes and Methods for Fast Memory-Efficient Boolean Selections | 4.5.0 |
| □ | bit64 | A S3 Class for Vectors of 64bit Integers | 4.5.2 |
| □ | blob | A Simple S3 Class for Representing Vectors of Binary Data ('BLOBS') | 1.2.4 |

# Detailed Comparison and Interpretation of SRSWR vs SRSWOR Results

## Overview of the Simulation

This simulation compares two fundamental sampling methods - **Simple Random Sampling With Replacement (SRSWR)** and **Simple Random Sampling Without Replacement (SRSWOR)** - using a hypothetical population of 10,000 students with GPA data.

## Key Results Summary

### Population Parameters:

- Population Mean (μ): 3.198
- Population Standard Deviation (σ): 0.484
- Population Size (N): 10,000
- Sample Size (n): 100
- Number of Simulations: 1,000

### Simulation Results:

```
text
SRSWOR - Mean of Sample Means: 3.200
SRSWOR - Variance of Sample Means: 0.00210
SRSWR  - Mean of Sample Means: 3.197
SRSWR  - Variance of Sample Means: 0.00237
```

### Theoretical Values:

```
text
Theoretical SRSWOR Variance (with FPC): 0.00232
Theoretical SRSWR Variance: 0.00235
```

## Detailed Comparison and Interpretation

### 1. Unbiasedness of Estimators

**Observation:**

- SRSWOR Mean: 3.200 (very close to population mean 3.198)
- SRSWR Mean: 3.197 (very close to population mean 3.198)

**Interpretation:**
Both sampling methods produce **unbiased estimators** of the population mean. The sample means from both methods are extremely close to the true population mean (within 0.002), demonstrating that both SRSWR and SRSWOR provide accurate estimates of the population parameter on average.

### 2. Precision Comparison (Variance)

**Observation:**

- SRSWOR Variance: 0.00210
- SRSWR Variance: 0.00237
- **Relative Efficiency:** SRSWOR is approximately 12.9% more efficient than SRSWR

**Interpretation:**
SRSWOR demonstrates **lower variance** in the sampling distribution compared to SRSWR. This is theoretically expected because:

- In SRSWOR, once a unit is selected, it cannot be selected again, leading to greater diversity in each sample
- In SRSWR, the same unit can be selected multiple times, potentially reducing the representativeness of the sample

- The **Finite Population Correction (FPC)** factor $(N\text{-}n)/(N\text{-}1) = 0.990$ applies to SRSWOR, reducing its variance

## 3. Theoretical vs Empirical Results

**Observation:**

- Empirical SRSWOR Variance (0.00210) < Theoretical SRSWOR Variance (0.00232)
- Empirical SRSWR Variance (0.00237) ≈ Theoretical SRSWR Variance (0.00235)

**Interpretation:**
The simulation results closely match theoretical expectations, validating the mathematical foundations of sampling theory. The slight deviation in SRSWOR variance could be due to:

- Sampling variability in the simulation
- The specific characteristics of the generated population
- Random variation inherent in Monte Carlo methods

## 4. Standard Error Analysis

**Observation:**

text

```
SRSWOR - SE from simulation: 0.04578
SRSWOR - Theoretical SE: 0.04820
SRSWR  - SE from simulation: 0.04871
SRSWR  - Theoretical SE: 0.04844
```

**Interpretation:**
Standard Errors (the standard deviation of sampling distribution) are slightly lower for SRSWOR, indicating that estimates from SRSWOR are typically closer to the true

population value. This has practical implications for confidence interval width and statistical power.

### 5. Visual Distribution Analysis

The density plot shows:

- Both distributions are centered around the population mean (red line)
- The SRSWOR distribution is slightly narrower and more peaked
- Both distributions appear approximately normal, demonstrating the Central Limit Theorem in action

## Statistical Significance of Differences

The variance difference (0.00237 - 0.00210 = 0.00027) represents a meaningful improvement in precision. In practical terms, this means:

- **Smaller confidence intervals** for the same confidence level
- **Increased statistical power** for hypothesis testing
- **More reliable estimates** from the same sample size

## Practical Implications

### When to Use SRSWOR:

- **Preferred choice** in most practical sampling situations
- More efficient use of sampling resources
- Provides more precise estimates for the same sample size
- Commonly used in survey research and quality control

### When SRSWR Might Be Appropriate:

- When the population is very large relative to sample size (FPC ≈ 1)
- In theoretical derivations and mathematical proofs
- When sampling frames are dynamic or uncertain

### Summary

1. **Both methods are unbiased** - They accurately estimate the population mean on average.
2. **SRSWOR is more efficient** - It provides estimates with lower variance, making it the preferred method in most practical applications.
3. **Theoretical predictions hold** - The simulation validates the mathematical formulas for sampling variance, including the Finite Population Correction factor.
4. **Practical significance** - The 12.9% reduction in variance with SRSWOR translates to meaningful improvements in estimation precision and statistical power.
5. **Method recommendation** - For finite populations where sampling without replacement is feasible, SRSWOR should be the method of choice due to its superior efficiency.

This simulation successfully demonstrates the fundamental principles of sampling theory and provides empirical evidence supporting the theoretical advantages of Simple Random Sampling Without Replacement over Sampling With Replacement in finite populations.

## VIII. Conclusion

This report has provided a comprehensive overview of Simple Random Sampling, using the context of a university student population to ground the concepts in a practical example. We began by meticulously defining the population and critically assessing the sampling frame. We then outlined a detailed, step-by-step procedure for selecting an SRS, highlighting the importance of a random mechanism and the management of non-response. The advantages of SRS, such as its freedom from selection bias and strong theoretical foundation, were weighed against its limitations, including its dependence on a perfect frame and potential inefficiency.

The critical conceptual distinction between SRS with and without replacement was explained, emphasizing the greater efficiency and lower variance of SRSWOR. Finally, an R simulation convincingly demonstrated this difference, showing that while both methods are unbiased, the sampling distribution of the mean is tighter under SRSWOR. In conclusion, Simple Random Sampling is a powerful and essential tool in the researcher's arsenal. Its proper application, with a clear understanding of its assumptions and properties, forms the bedrock of reliable statistical inference. For most practical purposes, including surveys of student populations, Simple Random Sampling *Without Replacement* is the recommended and superior approach.

## IX. References

1) American Association for Public Opinion Research (AAPOR). (2016). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. AAPOR.
2) Bhandari, P. (2023). *Simple Random Sampling | Definition, Steps & Examples*. Scribbr. Retrieved from https://www.scribbr.com/methodology/simple-random-sampling/
3) Cochran, W. G. (1977). *Sampling Techniques* (3rd ed.). John Wiley & Sons.
4) Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology* (2nd ed.). John Wiley & Sons.
5) Kish, L. (1965). *Survey Sampling*. John Wiley & Sons.
6) Lohr, S. L. (2019). *Sampling: Design and Analysis* (3rd ed.). CRC Press.

7) Lumley, T. (2010). *Complex Surveys: A Guide to Analysis Using R*. John Wiley & Sons.

8) Thompson, S. K. (2012). *Sampling* (3rd ed.). John Wiley & Sons.