Journal of University Studies for inclusive Research (USRIJ)
مجلة الدراسات الجامعية للبحوث الشاملة
ISSN: 2707-7675

Journal of University Studies for Inclusive Research
Vol.5, Issue 49 (2026), 157899- 157917
USRIJ Pvt. Ltd

Title: An Adaptive Ranked Set Sampling Framework for Mean Estimation under the Gompertz Distribution: A Simulation-Based Analysis

Author:

Sereen Ghazi Ahmad Al-Qwasmi

Affiliation:

Department of Statistics

Yarmouk University – Jordan

Email: qwasmisereen@gmail.com

Corresponding Author: Sereen Ghazi Ahmad Al-Qwasmi

Abstract

In this research, we developed a new method to calculate the average in the Gompertz distribution using AI-RSS, which is a method based on the intelligent and adaptive arrangement of samples that combines Simple Random Sampling and Ranked Set Sampling. We simulated the data using R to compare the performance of three methods: Simple Random Sampling (SRS), Ranked Set Sampling (RSS) and Artificial Intelligence Ranked Set Sampling (AI-RSS). The results showed that the smart method is better than other methods and gave more accurate averages, lower MSE and higher relative efficiency exceeding (2.8) compared to the random

sample. This shows that the integration of traditional statistical concepts with modern techniques gives more accurate and more reliable estimation methods.

Keywords: Gompertz distribution, adaptive sampling, AI-RSS method, mean estimation, simulation study, relative efficiency.

1. Introduction

It is very important to know how to calculate the average of a set of numbers. The average helps us to understand what is happening to the data (Cochran, 1977). A simple random sampling is usually used for its ease, but sometimes it does not give an accurate picture, for example, there is a large garden with many plants, in which case measuring the plant will be very difficult and stressful. Previously Mclntyre (1952) proposed a new idea and the way it works is as follows: we choose small simpots first, organise them quickly, and choose a unit from each group, and this method often gives a close mean of reality (Dell & Clutter, 1972; Ozturk, 2005). Some researchers have improved the method so that they repeated the order or used additional information to make the choices accurate (Mahdizadeh & Zamanzade, 2018). Recently, the use of machine learning has become useful. For example, Random Forest can guess which items are higher and which are lower. This helps us to choose samples in a smart way. We can change our choice according to the data (Breiman, 2001; Yu & Kumbier, 2020). The distribution of Gumbertiz is very visible in growth and survival studies, the numbers are often skewed and the possibilities are desired over time, the use of the order of samples helps to obtain the best estimate of the average and the addition of the machine increases the accuracy (Mazucheli et al., 2019; Gul & Yeniay, 2025). In this research, the Simple Random Sampling (SRS), the traditional sample order and the AI-RSS smart version will be compared, and simulated data will be used to see which method gives us an average closer to reality. The goal in this study is to find an easy and accurate way to calculate the average when the data is irregular.

Journal of University Studies for inclusive Research (USRIJ)
مجلة الدراسات الجامعية للبحوث الشاملة
ISSN: 2707-7675

2. Literature review

In the 20th century, the order of RSS samples became important in statistics, where new suggestions were made to help us take samples in an easy and effective way, especially when it is difficult to measure everything in society (Mclntyre, 1952). The idea is as follows, which is to divide the community into small groups and arrange the elements quickly according to the values that we see or can observe, and then choose a unit from each order to form the final sample. This method usually gives a more accurate average compared to the Simple Random Sampling (SRS) without what we need to provide in the sample size (Mclntyre, 1952). Over time, the order of samples became used in many fields such as environment, social sciences, agriculture, economics, and medical statistics (Chen, Bai, & Sinha, 2004). In a previous study, it was shown that RSS is better than SRS because you want to estimate the average weight of trees because the method more accurately represents society. It is considered an important step in the development of RSS because it gives high efficiency when society is diverse (Cochran, 1977). Some researchers have tried to develop a method to reduce errors resulting from manual arrangement. The most prominent of these improvements is Repeated RSS and Stratified RSS, which aims to reduce variation and increase the accuracy of measurements (Ozturk, 2011). The accuracy of the RSS depends on the quality of the initial order, as some studies have focussed on the effect of errors in the order of the elements. Incorrectly ordering items reduces efficiency and that using algorithms based on probability rules instead of manual reporting can minimise this effect and mitigate experimental bias (Mahdizadeh & Zamanzade, 2018). In the last decade, there has been a great interest in integrating machine learning and artificial intelligence with sampling methods. The use of machine learning algorithms helps to automatically organise items based on their statistical properties and this improves the efficiency of selecting representative samples (Goga, 2024). Neural Network and Random Forest algorithms help improve sample order and estimate

Journal of University Studies for inclusive Research (USRIJ)
مجلة الدراسات الجامعية للبحوث الشاملة
ISSN: 2707-7675

accuracy (Breiman, 2001). As for the Gompertz distribution, it is a flexible model for showing data that is deviated to the top, especially in studies of survival and biological growth. The distribution shows phenomena in which the risk rate increases over time, such as the average age or growth of organisms (Mazucheli et al., 2019). Several studies comparing RSS and SRS to estimate Gompertz coefficients using simulation have shown that RSS is better at minimising MSE and increasing relative efficiency RE. In short, the integration of artificial intelligence with the design of ordered samples represents a modern and promising trend in applied statistics. Although the studies dealt with RSS and many modifications to it, research that combines the distribution of Gompertz and machine learning algorithms within the AI-RSS design is still limited. This demonstrates the importance of this research as an innovative contribution because it combines the classic method and machine intelligence to give more accurate and flexible estimates.

3. Methodology

3.1. Study Design

I want to develop a new smart method of sampling called AI-RSS, in order to accurately calculate the average in a statistical community that follows the distribution of Gompertz. And compare its performance with two traditional methods: Simple Random Sampling (SRS) and the Ranked Set Sampling (RSS). Statistical simulations were used in research and a homogeneous artificial society was built before experiments in order to compare all methods with the same conditions and calculate performance indicators such as bias, RMSE, MSE, and relative efficiency RE.

3.2. The statistical community distribution of Gompertz

USRIJ
Journal of University Studies for inclusive Research (USRIJ)
مجلة الدراسات الجامعية للبحوث الشاملة
ISSN: 2707-7675

The statistical community is represented using the Gompertz distribution because it is flexible and estimated to represent the upward-deviant data, this distribution is often used in survival studies, population growth and biological modelling.

3.3. Sampling methods

3.3.1. Simple Random Sample (SRS): *n=50* units of the community were randomly selected without repetition. This method is the main reference for comparing the performance of other methods.

3.3.2. TRanked Set Sampling (RSS):

1. The community was divided into small groups, each group has 5 elements.
2. Each group was first arranged using a visual or technical assessment and added some small random noise to mimic the imperfect arrangement.
3. I chose one unit of each order to make the final sample.
4. The goal of this method is to minimise variation and improve the accuracy of the average estimate compared to the SRS.

3.3.3. Artificial Intelligence Ranked Set Sampling (AI-RSS):

1. The Z-coupler and Random Forest model were used to predict the order of items within each group.
2. The sample was selected adaptively according to the predictions and the model is retrained after a specific number of measurements to improve the accuracy of the arrangement and minimise bias.

Journal of University Studies for inclusive Research (USRIJ)
مجلة الدراسات الجامعية للبحوث الشاملة
ISSN: 2707-7675

3. This method combines the traditional RSS principle with artificial intelligence technologies and is very effective, especially with distorted or very variable data.

3.4. Description of the AI-RSS algorithm (process steps)

1. **Pilot sample:** We start by choosing a small initial sample from the community of *(n=50)* units so that we can train the initial prediction model, and the idea is to take samples that represent the community in an initial way before we begin to adapt.
2. **Auxiliary variable:** We create an auxiliary variable $\boldsymbol{z = X + \epsilon}$, associated with the original variable X but in a little noise and the goal is to simulate natural errors in measurement as if we were measuring real life instead of ideal data.
3. **Predictive model:** We train the **Random Forest** model using the plug variable Z to predict the values of X. This model will help us organise the items accurately before selecting the sample.
4. **Group composition:** We evaluate samples for small groups, each group has **5** elements according to the traditional Ranked Set Sampling method.
5. **Predictive order:** We arrange the elements of a group according to the values predicted by the **Random Forest** model, not by the real values, and this minimises errors in the selection of important elements.
6. **Selection**: We choose the item that corresponds to the required rank in the RSS course, which means we take the best item in the order expected from the model.
7. **Adaptive updating:** After every **5** new measurements, we return to train the **Random Forest** model by adding new observations, and this is how the order is updated and reduces the alignment with each course.
8. **Mean estimation:** After collecting samples for each method (SRS, RSS, AI-RSS), we calculate the average in a simple way: we add all the values taken from the sample and divide

Journal of University Studies for inclusive Research (USRIJ)
مجلة الدراسات الجامعية للبحوث الشاملة
ISSN: 2707-7675

them by their number, and the final average is the value that represents the data centre for each method and will be the basis for calculating the rest of the indicators such as bias, the average of the error boxes and the square root of the error boxes. The AI-RSS average is calculated from the last 50 measurements collected after each adaptive cycle because it reflects the best estimate of the average after updating the predictive model and reducing errors.

3.5 The justification for choosing the sample size n=50

We chose the sample size *(n=50)* because our main goal is to compare the performance between the different methods of all methods, but the order of relative performance between them will not change.

3.6. Definition of Relative Efficiency

To calculate the relative efficiency of each method, we used the formula:

$$RE = \frac{MSE_{SRS}}{MSE_{method}}$$

Means simply comparing the average error box for each method with the simple random sample to see which method has given a greater improvement in accuracy.

3.7. Simulation procedures

- 1000 experiments (nsim=1000) were performed per sampling method.
- In each experiment, the estimated average of the sample was calculated.

Journal of University Studies for inclusive Research (USRIJ)
مجلة الدراسات الجامعية للبحوث الشاملة
ISSN: 2707-7675

- The same community standards were used in every way to ensure a fair comparison.

3.8. Performance indicators

The performance of each method was evaluated using:

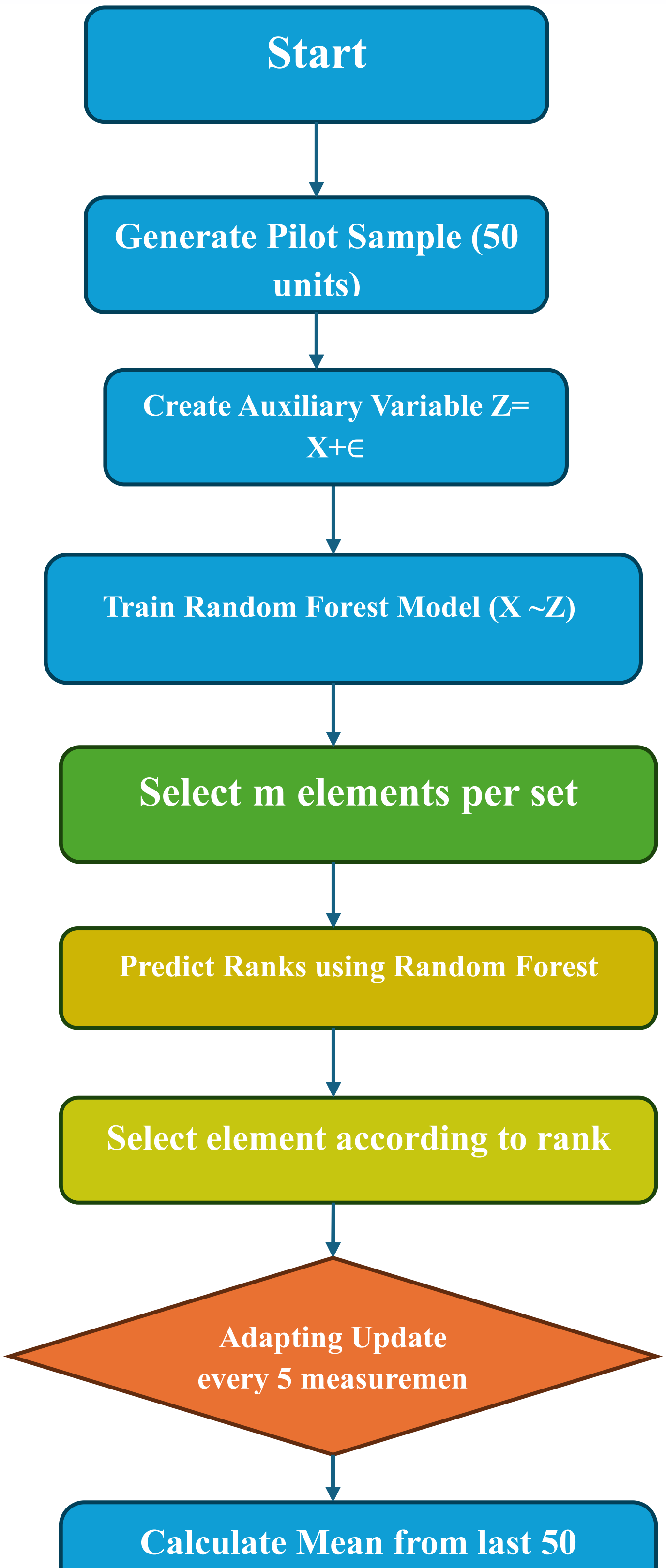
1. **Bias:** The difference between the estimated average and the real average.
2. **Square Average (MSE):** The average of the squares of the difference between the estimated and the real average.
3. **Variance:** Variation of averages across all experiments.
4. **Square root of the average squares (RMSE):** to facilitate the interpretation of the results in data units.
5. **Relative efficiency (RE):** Compare the performance of the developed methods with the SRS to show improvement in accuracy.

3.9. Analysing Tools

Simulated using R Studio and based on specialised simulation packages and Random Forest models, the code is designed so that the experiment can be easily repeated, full control of the characteristics of the community and accurate calculation of all statistical indicators.

Journal of University Studies for inclusive Research (USRIJ)
مجلة الدراسات الجامعية للبحوث الشاملة
ISSN: 2707-7675

Journal of University Studies for inclusive Research (USRIJ)
مجلة الدراسات الجامعية للبحوث الشاملة
ISSN: 2707-7675



Journal of University Studies for inclusive Research (USRIJ)
مجلة الدراسات الجامعية للبحوث الشاملة
ISSN: 2707-7675

4. Results

In this research, I decided to present the results of estimating the average in a statistical community that follows the Gompertz distribution using three methods:

- **Simple Ranked Sampling (SRS).**
- **Ranked Set Sampling (RSS).**
- **Artificial Intelligence Ranked Set Sampling (AI-RSS).**

A thousand experiments per method were performed on a community of (10,000) units with a sample size of *(n=50)* units per experiment, and an order set of 5 elements, to compare performance using indicators of bias, MSE, variance, RMSE, and relative efficiency compared to SRS.

4.7. Statistical findings

Table (1): Comparison of Sampling Methods under the Gompertz Distribution

Method	Bais	MSE	Variance	RMSE	Mean Est	n	RE vs SRS
AI-RSS (adaptive)	-0.0018	0.00508	0.00509	0.0713	1.1903	1000	2.86
RSS (imperfect)	-0.0010	0.00619	0.00620	0.0787	1.1911	1000	2.35
SRS	0.0016	0.01456	0.01457	0.1207	1.1937	1000	1.00

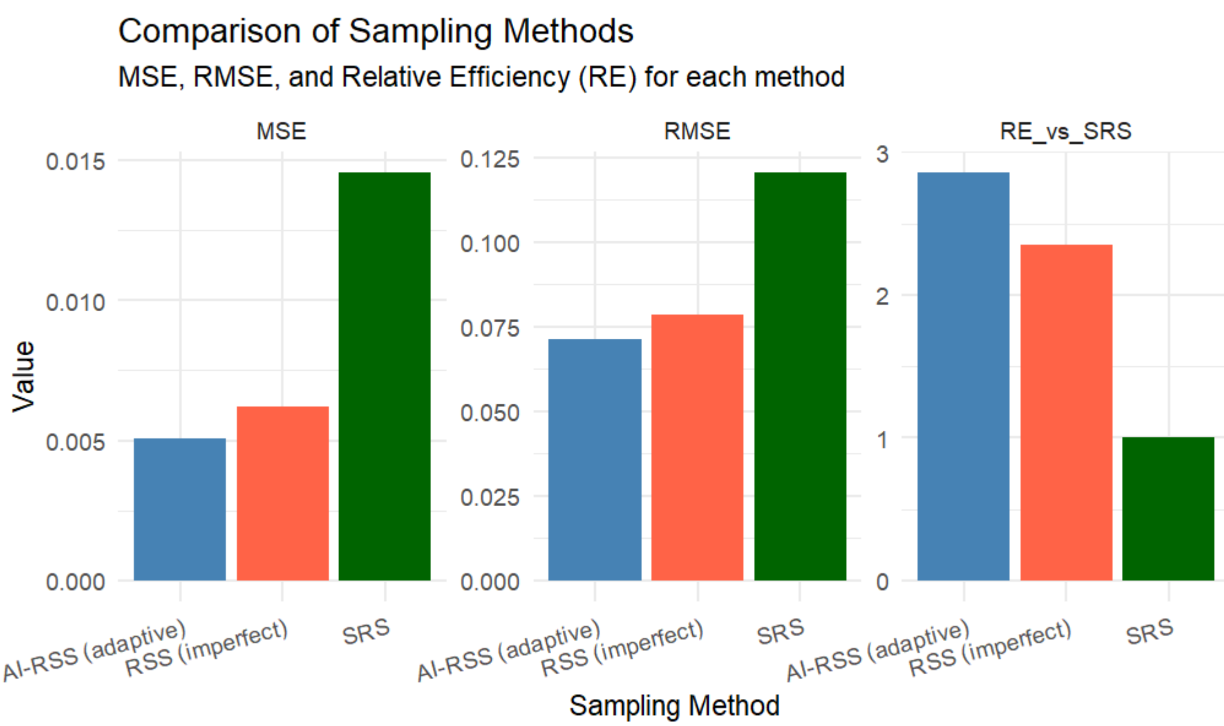
**Note: All analyses were done using R Version 4.5.1 on Windows to ensure that results can be replicated.*

**The figure 1 graph also shows the comparison of the performance of the three methods based on MSE, RMSE, and relative efficiency.*

Journal of University Studies for inclusive Research (USRIJ)
مجلة الدراسات الجامعية للبحوث الشاملة
ISSN: 2707-7675

Table (1) shows the difference between the three methods in the average estimate, the SRS method was the weakest in terms of accuracy, showing the highest MSE values and the highest value of Variance, while the traditional RSS method improved the results slightly but did not reach the AI-RSS level. The AI-RSS intelligent adaptive method achieved the lowest error and the highest relative efficiency, which means that it was the most stable and closest to the true value compared to the other two methods.

Figure (1): Performance comparison of SRS, RSS, and AI-RSS methods based on MSE, RMSE, and RE



**Note: All analyses were done using R Version 4.5.1 on Windows to ensure that results can be replicated.*

Journal of University Studies for inclusive Research (USRIJ)
مجلة الدراسات الجامعية للبحوث الشاملة
ISSN: 2707-7675

**The figure 1 graph also shows the comparison of the performance of the three methods based on MSE, RMSE, and relative efficiency.*

Figure (1) shows how the three methods clearly differ in terms of MSE error level, the SRS method remained the highest in RMSE, which means that it was the least accurate, and the RSS method was better than it but still medium, while the AI-RSS method was the least wrong and the most efficient, and this shows the effect of smart ranking and machine learning in improving the average estimate.

4.8. Data Analysis

4.8.1. Simple Random Sampling (SRS): It showed the highest values for MSE and RMSE among the three methods. This reflects the limits of the randomised sample when the data are very variable. However, bias was low, which meant that the estimates were unbiased but less efficient than guided methods.

4.8.2. Ranked Set Sampling (RSS): MSE and RMSE values were significantly reduced compared to the SRS while maintaining a low bias, and the relative efficiency was 2.35, which means that the order of the samples even if it is not perfect, it gives a significant improvement with the accuracy of the estimates.

4.8.3. Artificial Intelligence Ranked Set Sampling (AI-RSS):

- Showed the best performance among the three methods, with the lowest values of bias, MSE and RMSE and the highest relative efficiency RE=2.86.

- Using Random Forest to anticipate the order of items within groups helped to select samples that are more representative of the community and this reduced the error and increased the reliability of estimates.
- This shows that integrating artificial intelligence with the ordering of samples is an innovative and effective method of designing statistical experiments.

4.9.Discussion of the results of analysis

1. The results are consistent with previous studies that confirmed the superiority of RSS over SRS in estimating averages, especially in deviant or high-variable communities.
2. The basic innovation here is the integration of machine learning algorithms with ordered samples and this improves the order of the elements and minimises experimental error.
3. The research shows that AI-RSS is not just an optimisation of traditional RSS but an integrated model capable of dealing with complex data and extensive applications in survival studies, biological data, economics, big data analysis, and environmental and agricultural models.

All methods showed low bias and this emphasises the fairness of estimates and the importance of choosing the appropriate method according to the nature of the statistical community.

5. Discussion and conclusions

5.1. Innovation and scientific significance

In this research, I tried to combine the traditional RSS method and AI-RSS techniques to create a more accurate and efficient sampling method, as the results showed that smart and adaptive

Journal of University Studies for inclusive Research (USRIJ)
مجلة الدراسات الجامعية للبحوث الشاملة
ISSN: 2707-7675

methods improve the arrangement of elements within groups and reduce experimental errors compared to traditional methods, especially when the data is very distorted or variable.

This research makes an innovative contribution because it:

- Design smart and adaptive samples.
- The use of ordered samples has expanded to include modern machine learning techniques.
- Improve the average estimate without the need to increase the sample size.

5.2. Discussing the results

1. AI-RSS outperforms traditional RSS and SRS demonstrates the practical value of AI in sample design.
2. Low bias in all ways this confirms that the estimates are unbiased, this reflects the higher relative efficiency of AI-RSS and this achieves a clear improvement in performance.
3. The results are consistent with previous studies on RSS effectiveness but go beyond it by adding adaptive predictability.
4. The research provides practical evidence that statistically intelligent designs are capable of handling complex data and this makes them suitable for use in recent lessons.

5.3. Future applications

1. Using AI-RSS for survival studies and biometric data to calculate averages with high accuracy.

Journal of University Studies for inclusive Research (USRIJ)
مجلة الدراسات الجامعية للبحوث الشاملة
ISSN: 2707-7675

2. Relying on artificial intelligence algorithms can replace manual ranking and minimise human errors.
3. Research opens the way for the application of machine learning in the design of samples for large and varied distributions of data.
4. In the future, we can expand the method to include non-Jumbertz distributions and use larger advanced algorithms such as neural networks.

5.4. Practical recommendations

1. The use of AI-RSS in applied studies that require an average estimate with high accuracy.
2. Integrate auxiliary variables and machine learning models to optimise the order of samples in complex communities.
3. Using the simulation to evaluate the performance of the method before applying it to real data.
4. The development of advanced AI-RSS frameworks that include modern machine learning algorithms to increase relative efficiency and improve accuracy.

5.5. Future business

1. The AI-RSS experiment on real data from different fields such as economic or biological studies to emphasise the strength of the method outside the simulation environment.
2. Testing the method on larger sample sizes and more diverse communities to determine the effect of sample size on average accuracy and relative efficiency.
3. Apply AI-RSS to different distributions such as normal or logarithmic distribution to expand the use of the method.

Journal of University Studies for inclusive Research (USRIJ)
مجلة الدراسات الجامعية للبحوث الشاملة
ISSN: 2707-7675

4. Try other machine learning algorithms such as neural networks to optimise sample arrangement and minimise errors.
5. Integrate additional assistive variables within the model to optimise performance when handling complex or multidimensional data.

5.6. Limitations

1. The search relied entirely on simulating data using R and was not applied to real data. This means that the results may vary slightly when dealing with real data containing noise or measurement errors.
2. The sample size was only 50 per experiment and this is a relatively small size compared to some applied studies and increasing the sample size can affect the accuracy of the estimates and the relative efficiency.
3. The study focussed only on the distribution of Jumbritz and the results may not apply to other distributions with different characteristics such as normalisation.
4. Ranking was done using Random Forest only and using other algorithms can give different results.

6. Conclusion

Research has shown that combining ordered samples with artificial intelligence gives a new accurate and reliable way to calculate the average in complex data. AI-RSS represents a step towards intelligent statistical designs that adapt to the nature of data and open a new horizon for applications in medicine, big data analysis and economics. This research demonstrates the importance of innovation in sample design and emphasises the possibility of significantly improving statistical performance without increasing resources or increasing the sample size.

Journal of University Studies for inclusive Research (USRIJ)
مجلة الدراسات الجامعية للبحوث الشاملة
ISSN: 2707-7675

The results indicate the importance of integrating artificial intelligence with sampling methods, especially in heterogeneous communities where the method of estimating parameters can change and increase their accuracy and reliability.

References

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chen, Z., Bai, Z., & Sinha, B. K. (2004). Ranked set sampling: Theory and applications. Springer.

Cochran, W. G. (1977). Sampling techniques (3rd ed.). John Wiley & Sons.

Dell, T. R., & Clutter, J. L. (1972). Ranked set sampling theory with an order-statistics background. Biometrics, 28, 545–555.

Goga, C. (2024). From traditional to modern machine-learning estimation methods for survey sampling. The Survey Statistician, 90, 18–29.

Gul, H. H., & Yeniay Köcer, N. (2025). Estimating the Gompertz distribution parameters under folded ranked set sampling. Sigma Journal of Engineering and Natural Sciences, 43(3), 737–747. https://doi.org/10.14744/sigma.2025.00068

Journal of University Studies for inclusive Research (USRIJ)
مجلة الدراسات الجامعية للبحوث الشاملة
ISSN: 2707-7675

Mahdizadeh, M., & Zamanzade, E. (2018). Smooth estimation of a reliability function in ranked set sampling. Journal of Theoretical and Applied Statistics, 52(4), 1–19. https://doi.org/10.1080/02331888.2018.1477157

McIntyre, G. A. (1952). A method for unbiased selective sampling using ranked sets. Australian Journal of Agricultural Research, 3(4), 385–390. https://doi.org/10.1071/AR9520385

Öztürk, Ö. (2005). Ranked set sampling: Theory and applications. Journal of the American Statistical Association, 100(471), 1093–1094. https://doi.org/10.1198/jasa.2005.s35

Singh, R., & Mangat, N. S. (2013). Elements of survey sampling. Springer.

Yu, B., & Kumbier, K. (2020). Veridical data science. Proceedings of the National Academy of Sciences, 117(8), 3920–3929. https://doi.org/10.1073/pnas.1901326117