Optimizing Sample Size for Accelerated Failure Time Model in Progressive Type-II Censoring through Rank Set Sampling

Ibrahim Alliu*, Lili Yu, Hani Samawi and Jing Kersey

Jiann-Ping Hsu College of Public Health, Georgia Southern University, USA

Abstract: Survival data is a type of data that measures the time from a defined starting point until the occurrence of a particular event, such as time to death from small cell lung cancer after diagnosis, Length of time in remission for leukemia patients, Length of stay (i.e., time until discharge) in hospital after surgery. The accelerated failure time (AFT) models are popular linear models for analyzing survival data. It provides a linear relationship between the log of the failure time and covariates that affect the expected failure time by contracting or expanding the time scale. This paper examines the performance of the Rank Set Sampling (RSS) on the AFT models for Progressive Type-II censoring-survival data. The Ranked Set Sampling (RSS) is a sampling scheme that selects a sample based on a baseline auxiliary variable for assessing survival time. Simulation studies show that this approach provides a more robust testing procedure, and a more efficient hazard ratio estimate than simple random sampling (SRS). The lung cancer survival data are used to demonstrate the method.

Keywords: Accelerated failure time model, Hazard ratio, Progressive Type-II censoring, Survival analysis.

INTRODUCTION

Survival analysis is a statistical tool used to study the time until a subject experiences an event of interest. This statistical method is widely used in various fields, including medicine, social science, engineering, and finance. Survival data are time-to-event data consisting of distinct start and end times. Survival data include the time between surgery and death, the time between treatment and the appearance of another disease (e.g., a tumor), and the time between response and recurrence. In real-world scenarios, survival data often follows a Weibull distribution, commonly used to model the time to failure of various systems or components. However, life-testing experiments are challenging due to the difficulty of getting complete information on all failed units. As a time-saving measure and to reduce costs, the sampling process is truncated according to a predetermined censoring scheme.

Censoring refers to incomplete or partial information about the event. It happens when the event does not occur or is not observed during the study period. There are several reasons why censoring is common in survival analyses, such as the study having a fixed endpoint, some participants not being followed up, or the event not occurring for all participants. A relatively effective censoring scheme design can simultaneously achieve cost savings and efficient statistical inference.

*Address correspondence to this author at the Jiann-Ping Hsu College of Public Health, Georgia Southern University, 501 Forest Dr, Statesboro, GA 30458, USA; E-mail: ia00702@georgiasouthern.edu

The most conventional censoring schemes are Type-I and Type-II censoring. In type I, the experiment is terminated at a predetermined time, whereas in type II, the experiment is terminated at a predetermined number of failures. However, removing active units during the experiment is prohibited in both types.

In some cases, it may be desirable to remove items being tested before their predetermined termination points, intentionally or unintentionally, to reduce the cost of the experiment and the time consumed. An example is the study of the weariness of units that must be worn entirely or disintegrated at different experiment stages during their aging process, which is quite time-consuming. Another example is the early removal of some surviving units in the experiment so they can be used in other experiments to minimize the cost of the experiment.

This leads to the practice of Progressively Type II censoring, which many researchers consider a practical approach to minimizing the cost and the time consumed. It incorporates ordinary order statistics (OS) and type II censoring, making it highly desirable and used in experimental design.

There have been a vast number of discussions on progressive censoring and its applications; interested readers may refer to the books by Balakrishnan & Aggarwala and Balakrishnan & Cramer [1] for recent reviews and discussions of the need for this type of censoring.

Progressive censoring can be conducted as follows. Suppose n identical subjects are put through a lifetime test, and m failures are observed ($n > m$). After the first failure, R_1 surviving items are immediately removed, leaving $(n-1-R_1)$ items still in the test. After the second failure, R_2 more surviving items are randomly selected and removed. The process continues until we observe m failures. At this point, all the remaining $n-m-R_1-\cdots-R_{m-1}$ $(=R_m)$ surviving subjects are removed from the experiment. We assume that these *subjects'* lifetimes are independent and identically distributed. The values of *n, m,* and the number of removals at each stage R_1, *{i=1,2, . . . , m} all have predetermined values[2]*.

If $R_1=R_2=\cdots=R_{m-1}=0$, so $R_m=n-m$, which corresponds to Type-II censoring. If $R_1=R_2=\cdots=R_m=0$ so, then (*m* = *n*) represents the complete data set. Check Balakrishnan and Cramer [3] for a comprehensive literature review on progressive censoring.

The progressive Type-II censoring scheme has several advantages over other types of censoring. One of the main advantages is that it can provide more information about the distribution of failure times than other censoring schemes. It can also reduce the number of test units required to obtain a given precision level in estimating the distribution parameters.

In survival analysis, selecting an efficient sampling method is crucial for obtaining estimates while minimizing resource use. One promising approach that enhances efficiency compared to simple random sampling (SRS) is ranked set sampling (RSS). This method involves ranking the units within a population using an auxiliary variable and selecting the sample for analysis.

Rank Set Sampling (RSS) is a data collection technique that preserves the fundamental properties of Simple Random Samples (SRS) while leveraging additional information available in the population[3]. By creating an "artificially stratified" sample with enhanced structure, RSS ensures a more structured selection process, leading to measurements that better represent the full range of values within the population.

Stepwise Process for RSS

1. Define k (set size) and m (number of cycles).
2. Identify an auxiliary variable (e.g., expert judgment, visual assessment, or another measurable characteristic) for ranking.
3. Draw a Simple Random Sample (SRS) of size k from the population.
4. Rank the k units based on the auxiliary variable (without measurement).
5. Select and Measure One Unit:
   - ❖ Choose the smallest ranked unit from the first sample → Denote as $X_{[1].}$
   - ❖ The remaining k - 1 unmeasured units are discarded.
6. Repeat for Subsequent Samples:
   - ❖ Draw another random sample of k units.
   - ❖ Rank the units again.
   - ❖ Select the second smallest unit as the next measured observation $X_{[2]}$.
7. Continue this process until the largest unit from the k-th sample is selected, denoted as X[k].
8. One RSS cycle results in k measured observations: X[1], X[2], …, X[k].

Repeat Steps 3-8 for m independent cycles to achieve the desired total sample size $n=km$. The final Ranked Set Sample consists of *n* measured units, ensuring an evenly distributed selection across rank positions, capturing a broader and more structured representation of the population compared to Simple Random Sampling (SRS).

While RSS is effective in increasing the representativeness of a sample, there has been no research on how the Weibull survival model behaves under Progressive Type-II censoring with RSS. This study aims to fill this gap by giving a complete framework for simulating progressive Type-II censored survival times with different censoring schemes and rates using a Ranked Set Sampling scheme.

2. PRELIMINARIES

Let *T* be a random variable representing the failure time of an event. Three key functions typically characterize the distribution of T:

1. Survival Function S(t): Represents the probability that the event has not yet occurred by time t.
2. Hazard Rate Function h(t): The risk function describes the instantaneous rate at which the event occurs, given that it has not happened before time t.
3. Probability Density Function (PDF) f(t) or Probability Mass Function (PMF) for discrete cases: Represents the likelihood of the event occurring at a specific time t.

In survival analysis, censoring occurs when the exact time of the event is not fully observed. We restrict to the setting of Progressive Type-II Censoring (PTII), which is a scheme where some of the units are removed from the observation while they have not yet experienced the event (often referred to as failure or death), allowing for better usage of an experimental study, while maintaining statistical reliability.

Under the PTII censoring for life-testing, suppose that $T_{1:m:n} < T_{2:m:n} <\ldots< T_{m:m:n}$ are the lifetimes of the completely observed unit to fail and $\boldsymbol{R} = (R_1, R_2, \ldots, R_m)$ represent the number of units withdrawn from the experiment at these failure times. If the failure times are from an absolute continuous population with cumulative distribution function *F(t)* and probability density function *f(t)*, the joint probability density function for the progressive censored times $T_{1:m:n} < T_{2:m:n} < \ldots < T_{m:m:n}$ is given by:

$$f_{T_{1:m::n}, \ldots, T_{m:m:n}}(t_1 \ldots t_2, \ldots, t_m) = c \prod_{i=1}^{m} f(t_i)[1 - F(t_i)]^{R_i}, \quad -\infty < t_1 <\ldots t_m < \infty$$

$$l(\beta) = c \prod_{i=1}^{m} f(t_i) S(t_i)^{R_i} \quad (1)$$

where

$c = n(n-1-R_1)(n-2-R_1-R_2) \ldots (n - \sum_{i=1}^{m}(R_i+1))$ and β is the set of parameters [4].

2.1. Accelerated Failure Time Models (AFT)

The Accelerated Failure Time (AFT) model is an approach in survival analysis that examines the relationship between survival time and explanatory variables. In this model, the logarithm of survival time serves as the response variable, capturing how covariates influence the time until an event occurs. The model also includes an error term, which is assumed to follow a specific probability distribution. The key assumption in the AFT model is that covariates affect survival time in a multiplicative (proportional) manner, meaning that changes in predictor variables either accelerate or decelerate the time to failure.

Alternatively, the AFT model can be expressed as a linear relationship between the logarithm of survival time and an error term, where the error term follows a specific distribution, such as exponential, Weibull, log-normal, or log-logistic. This formulation allows for flexibility in modeling different types of survival data and is particularly useful when proportional hazard assumptions are unsuitable.

The general log-linear representation of the AFT model for the i-th individual, according to Liu [2,5] and Samawi *et al*. [5], is given as

$$logT_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \sigma\varepsilon_i, \ i = 1,2,\ldots,m \quad (2)$$

$$S_i(t) = P(T_i \geq t_i)$$

$$S_i(t) = P(logT_i \geq logt_i) \quad (3)$$

$$S_i(t) = P\left(\varepsilon_i \geq \frac{logt_i - \beta_0 - w_i'\beta}{\sigma}\right), \quad 0 < t_i < \infty$$

where $w_i = (x_{i1}, \ldots, x_{ip})'$, represents the vector of the observed covariates for the *i*-th individual, and $\beta = (\beta_1, \beta_2, \ldots, \beta_p)$ denote the coefficients of the regression coefficient.

The hazard function for T_i at time *t* is given by:

$$h_i(t|w_i, \beta) = \frac{1}{t_i\sigma} h_0\left(\frac{logt_i - \beta_0 - w_i'\beta}{\sigma}\right), i = 1,2, \ldots m \quad (4)$$

Here $h_0(t)$ is the baseline hazard function at survival time *t*. The effect of the covariates $\{X_1, X_2, \ldots, X_p\}$on the hazard rate is assumed to have a multiplicative effect. Consequently, the predicted hazard function, given the covariate values $\{x_{i1}, x_{i2}, \ldots, x_{ip}\}$, is denoted as $\hat{h}\{x_{i1}, x_{i2}, \ldots, x_{ip}\}$

In this study, we examine the performance of the AFT model under Progressive Type-II censoring.

2.2. Weibull AFT Regression Model

The Weibull distribution is commonly used to model the lifetimes or failure times of specific devices, systems, or biological entities. This distribution is practical because of its flexible shape parameters that allow it to model different failure behaviors, including increasing failure rates that account for fatigue failures over time. Weibull plotting is a graphical method that helps verify the Weibull assumption and estimate the two key distribution parameters. This method is useful for analyzing complete failure time data, and Type-II censored data, where only the lowest or failure times from a sample of size n are observed.

In Type-II censoring, *n* items are simultaneously put on a common test until the first *r* failures occur $2 \leq r \leq n$. The requirement $r \geq 2$ is needed to ensure sufficient spread in the failure time data, allowing for a meaningful Weibull probability plot. If all *n* items are observed until failure $r = n$, the dataset is considered complete rather than censored.

The two-parameter Weibull distribution provides a reasonable model for describing the variability in the failure time data. If T represents the generic failure time, then the Weibull distribution function of T is given by

$$F_T(t) = P(T \leq t) = 1 - e^{-\left(\frac{t}{\lambda}\right)^{\delta}}, \quad for \ t \geq 0 \quad (5)$$

The parameter λ is called the scale parameter or characteristic life, while δ is the shape parameter.

The Weibull distribution function (W (λ, δ)) follows an extreme value distribution when $log(T_{(i)})$ is expressed as a function of the parameters. Let's assume T is distributed as W (λ, δ), where λ is the scale parameter, and δ is the shape parameter.

According to Samawi *et al.* [5], the hazard function is given by

$$h_i(t|w_i,\beta) = (\delta^*)^{-1} e^{\left(\frac{logt_i - w_i'\beta}{\delta^*}\right)}, \quad i = 1,2,\ldots,m \tag{6}$$

where $\delta^* = \frac{1}{\delta}$ and β is the regression coefficient vector. The survival function of T is

$$S_i(t) = e\left[-e^{\left(\frac{logt_i - w_i'\beta}{\delta^*}\right)}\right], \quad -\infty < log\, t_i < \infty \tag{7}$$

and the pdf

$$f(t_i) = (\delta^*)^{-1} exp\left[\frac{logt_i - w_i'\beta}{\delta^*} - e^{\left(\frac{logt_i - w_i'\beta}{\delta^*}\right)}\right], \quad -\infty < logt_i < \infty \tag{8}$$

base on

$$f_{T_{1:m:n}, \ldots, T_{m:m:n}}(t_1 \ldots t_2, \ldots, t_m) = c\prod_{i=1}^{m} f(t_i)[1 - F(t_i)]^{R_i}$$

$$l(\beta) = c\prod_{i=1}^{m} f(t_i) S(t_i)^{R_i} \; .$$

The likelihood function based on progressive type II can be written as

$$L(\beta) = log[\,l(\beta)] \propto \sum_{i=1}^{m}\left[(-\,log\,\delta^*) + \left(\frac{log\,t_i - w_i'\beta}{\delta^*}\right) - (R_i + 1)\,exp\left(\frac{log\,t_i - w_i'\beta}{\delta^*}\right)\right] \tag{9}$$

2.3. Maximum Likelihood Estimates of the Parameters

The Maximum Likelihood Estimate (MLE) of the *j*th covariate parameter can be obtained by solving

$$\frac{\partial L(\beta)}{\partial \beta_j} = \sum_{i-1}^{m}\left[\frac{x_{ij}}{\delta^*}\left\{1 - (R_i + 1)e^{\left(\frac{logt_i - w_i'\beta}{\delta^*}\right)}\right\}\right] = 0 \tag{10}$$

to derive the $\hat{\beta}_j'$ and the Fisher information matrix can be obtained by using the partial second derivative

$$-E\left[\left(\frac{\partial^2 L(\beta)}{\partial\beta_j \partial\beta_{j'}}\right)\right]_{(p+1)\times(p+1)} = E\left(\sum_{i-1}^{m}\left[\frac{x_{ij}x'_{ij'}}{\delta^*}(R_i + 1)e^{\left(\frac{logT_i - w_i'\beta}{\delta^*}\right)}\right]\right) \tag{11}$$

3. STRUCTURE OF A RANKED SET SAMPLE

In statistical sampling, efficiency is crucial to obtaining precise estimates while minimizing costs. While Simple Random Sampling (SRS) remains a widely used method, it does not always yield the most representative sample, especially when additional information about the population is available. Ranked Set Sampling (RSS) offers an improved alternative by incorporating a ranking mechanism before selection, enhancing the sample's representativeness and reducing estimation errors.

To better understand the advantages of RSS, we compare it with SRS of the same size by looking at a single cycle with perfect ranking and set size *k*. Here, the ranked set sample observations are also the respective order statistics. Let $X_1, \ldots, X_k$ denote an SRS of size *k* from a continuous population with probability density function (p.d.f.) $f(x)$ and cumulative distribution function (c.d.f.) $F(x)$ [6]. Let $X_1^*, \ldots, X_k^*$ be an RSS of size *k* obtained from *k* independent random samples of k units each.

In a Simple Random Sample (SRS), the *k*-selected observations are independent, and each of them represents a typical value from the population. However, there is no additional structure imposed on their relationship to one another. Letting $X_{(1)} \le X_{(2)} \le \cdots \le X_{(k)}$ be the order statistics associated with these SRS observations. The joint probability distribution function (p.d.f.) is given by

$$h_{SRS}(x_{(1)}, \ldots, x_{(k)}) = k!\prod_{i=1}^{k} f(x_{(i)}) I_{\{-\infty < x_{(1)} \le x_{(2)} \le \cdots \le x_{(k)} < \infty\}}(x_{(1)}, \ldots, x_{(k)}) \tag{12}$$

However, in Ranked Set Sampling (RSS), additional information and structure are incorporated by the ranking process before sample selection. Unlike in Simple Random Sampling (SRS), where the order statistics are dependent, the *k*-ranked observations in RSS, denoted as $X_1^*, \ldots, X_k^*$ are independent variables, each providing information about a different aspect of the population. The joint p.d.f is given by

$$h_{RSS}(x_{(1)}^*, \ldots, x_{(k)}^*) = \prod_{i=1}^{k} f_i(x_{(i)}^*) \tag{13}$$

where

$$f_i(x_{(i)}^*) = \frac{k!}{(i-1)!(k-i)!}\left[F(x_{(i)}^*)\right]^{i-1}\left[1 - F(x_{(i)}^*)\right]^{k-1} f(x_{(i)}^*) \tag{14}$$

is the p.d.f. for the *i*-th order statistic for an SRS of size *k* from the population with p.d.f. *f(x)* and c.d.f. *F(x).* This additional structure introduced by the judgment-based ranking and the independence of the resulting order statistics makes RSS-based estimation procedures more efficient than those based on SRS, even when the number of measured observations is the same.

4. SIMULATION STUDIES

In this section, we conduct a series of simulation experiments to evaluate the performance of the

Accelerate Failure Time (AFT) model under progressive type II censoring, comparing Rank Set sampling (RSS) when the available ranked auxiliary covariate is positively associated with our variable of interest with Simple Random Sampling (SRS). For the simulation, the primary objective is to assess the efficiency of RSS in parameter estimation by analyzing its bias, mean squared error (MSE), confidence interval coverage, and statistical power.

4.1. Simulation Setup

We investigate the efficiency of the AFT model under different sampling and censoring schemes, and we generate progressively Type-II censored survival data from a Weibull distribution. The Weibull AFT model generates a large replicate of progressive Type-II censoring samples. We considered $\sigma = 1$, and the covariates β_1 at 0, 0.05, and 0.1. We study the performance of testing the hypothesis of no factor effect after controlling the auxiliary covariate. The process was repeated 1000 times.

In our simulations, we evaluated the performance of the model by calculating the Bias, MSE, test power $H_0: \beta_1 = 0\ vs.\ H_1:\ \beta_1 \neq 0$, and 95% confidence interval coverage. To assess the effect of censoring intensity, we considered a *n* = 300 to examine nine different failure rates, ranging from 60% to 96.7%.

The failure rates were defined as:

$$\left\{\frac{m}{n}\times 100\%\right\} = \{60,\ 66.7,\ 73.3,\ 76.7,\ 80,\ 83.3,\ 86.7,\ 93.3,\ 96.7\}.$$

Where the number of observed failures (m) corresponds to *m* = {180, 200, 220, 230, 240, 250, 260, 280, 290}, these failure rates allow us to analyze the impact of varying censoring levels on some of the efficiency of Rank Set Sampling (RSS) compared to Simple Random Samling (SRS).

4.2. Censoring Scheme Considered

We examine three different Progressive Type -II censoring schemes to examine how RSS performs under various conditions:

1. Left Censoring: The first (*n* - *m*) subjects are removed after the first failure.
2. Right Censoring: The last (*n* - *m*) subjects are removed after the *m*th failure.
3. Uniform Censoring: Censored subjects are removed evenly across all failure events.

Table **1** summarizes the censoring percentages used in the simulation.

4.3. Simulation Results

Due to space limitations, we presented only a subset of the simulation results. Table **2** illustrates the statistical power of testing the hypothesis:

$$H_0:\ \beta_1 = 0\ vs.\ H_1:\ \beta_1 \neq 0$$

when controlling for the ranked auxiliary covariate. The Ranked Set Sampling (RSS) results yield a more powerful test than Simple Random Sampling (SRS). Furthermore, we observe that the power of the test increases as the set size *m* grows, demonstrating the efficiency of RSS in detecting true effects.

Similarly, Table **3** presents a comparison of Mean Squared Errors (MSEs) and confidence interval widths for hazard ratio estimates. The findings confirm that RSS consistently produces smaller MSE values and narrower confidence intervals than SRS, highlighting its superiority in estimation accuracy.

We used the Ranked Set Sampling (RSS) method to estimate conditional hazard ratios from simulated data and plotted the Root Mean Squared Error (RMSE) and Bias in Figure **1** and Figure **2**, respectively. In

Table 1: Censoring Schemes and Percentages

% of Censored Data	(n, m)	Censoring Left	Censoring Right	Censoring Uniformly
40%	(300, 180)	(120, 0, . . . , 0)	(0, 0, . . . , 120)	(0,1,1,0,…,0,1,1)
33.3%	(300, 200)	(100, 0, . . . , 0)	(0, 0, . . . , 100)	(1,0,1,0,…,1,0)
26.7%	(300, 220)	(80, 0, . . . , 0)	(0, 0, . . . , 80)	(0,1,1,0,…,0,1,1)
23.3%	(300, 230)	(70, 0, . . . , 0)	(0, 0, . . . , 70)	$(7,0^{*22},\ldots,7,0^{*22})$
20%	(300, 240)	(60, 0, . . . , 0)	(0, 0, . . . , 60)	(0,1,1,0,…,0,1,1)
16.7%	(300, 250)	(50, 0, . . . , 0)	(0, 0, . . . , 50)	$(1,0^{*4},\ldots,1,0^{*4})$
13.3%	(300, 260)	(40, 0, . . . , 0)	(0, 0, . . . , 40)	(0,1,1,0,…,0,1,1)
6.7%	(300, 280)	(20, 0, . . . , 0)	(0, 0, . . . , 20)	$(1,0^{*13},\ldots,1,0^{*13})$
3.3%	(300, 290)	(10, 0, . . . , 0)	(0, 0, . . . , 10)	(0,1,1,0,…,0,1,1)

Table 2: Estimating the Power of Testing *HR* = 1 when α = 0.05

β_1	m	Left		Right		Uniform	
		SRS	RSS	SRS	RSS	SRS	RSS
0	180	0.055	0.0195	0.104	0.068	0.062	0.0265
	200	0.0565	0.0155	0.104	0.0405	0.065	0.014
	220	0.0555	0.016	0.0955	0.0415	0.0685	0.0185
	230	0.0555	0.0145	0.0875	0.0355	0.056	0.0125
	240	0.056	0.009	0.0925	0.0365	0.0625	0.014
	250	0.057	0.013	0.0905	0.031	0.0615	0.013
	260	0.06	0.0085	0.092	0.024	0.0675	0.011
	280	0.06	0.0115	0.074	0.0185	0.06	0.0075
	290	0.0535	0.0095	0.068	0.0155	0.054	0.0105
0.05	180	0.2495	0.172	0.5885	0.509	0.4135	0.3095
	200	0.283	0.2155	0.5755	0.531	0.4235	0.347
	220	0.293	0.2295	0.5625	0.511	0.4245	0.3585
	230	0.2975	0.254	0.5465	0.504	0.381	0.319
	240	0.3345	0.216	0.5355	0.439	0.406	0.2705
	250	0.333	0.2715	0.5135	0.4675	0.3985	0.336
	260	0.361	0.283	0.504	0.4415	0.4055	0.3255
	280	0.3765	0.2845	0.472	0.3915	0.398	0.335
	290	0.3965	0.3175	0.444	0.37	0.404	0.3205
0.1	180	0.7485	0.7435	0.978	0.9835	0.9325	0.943
	200	0.794	0.837	0.9765	0.992	0.9295	0.964
	220	0.822	0.87	0.9765	0.9875	0.931	0.9635
	230	0.8345	0.8915	0.972	0.988	0.9055	0.9465
	240	0.855	0.8835	0.9675	0.984	0.9265	0.9525
	250	0.8655	0.923	0.961	0.99	0.924	0.9585
	260	0.886	0.933	0.955	0.987	0.9275	0.9595
	280	0.907	0.941	0.955	0.9815	0.921	0.957
	290	0.9175	0.964	0.943	0.9785	0.924	0.967

Table 3: Estimating a 95% Confidence Interval of the Hazard Ratio (HR)

β_1	m	Left		Right		Uniform	
		SRS	RSS	SRS	RSS	SRS	RSS
0	180	0.945	0.9805	0.896	0.932	0.938	0.9735
	200	0.9435	0.9845	0.896	0.9595	0.935	0.986
	220	0.9445	0.984	0.9045	0.9585	0.9315	0.9815
	230	0.9445	0.9855	0.9125	0.9645	0.944	0.9875
	240	0.944	0.991	0.9075	0.9635	0.9375	0.986
	250	0.943	0.987	0.9095	0.969	0.9385	0.987
	260	0.94	0.9915	0.908	0.976	0.9325	0.989
	280	0.94	0.9885	0.926	0.9815	0.94	0.9925
	290	0.9465	0.9905	0.932	0.9845	0.946	0.9895

(Table 3). Continued.

β_1	m	Left		Right		Uniform	
		SRS	RSS	SRS	RSS	SRS	RSS
0.05	180	0.945	0.9835	0.896	0.937	0.938	0.973
	200	0.9435	0.9875	0.896	0.953	0.935	0.9785
	220	0.9445	0.983	0.9045	0.9625	0.9315	0.9835
	230	0.9445	0.991	0.9125	0.9635	0.944	0.9815
	240	0.944	0.988	0.9075	0.963	0.9375	0.9825
	250	0.943	0.987	0.9095	0.975	0.9385	0.9865
	260	0.94	0.989	0.908	0.9785	0.9325	0.988
	280	0.94	0.9891	0.926	0.9825	0.94	0.986
	290	0.9465	0.9905	0.932	0.99	0.946	0.992
0.1	180	0.945	0.9825	0.896	0.9355	0.938	0.976
	200	0.9435	0.99	0.896	0.9525	0.935	0.9815
	220	0.9445	0.9875	0.9045	0.955	0.9315	0.9785
	230	0.9445	0.9845	0.9125	0.9605	0.944	0.98
	240	0.944	0.9855	0.9075	0.969	0.9375	0.9855
	250	0.943	0.9905	0.9095	0.978	0.9385	0.987
	260	0.94	0.991	0.908	0.9825	0.932	0.9815
	280	0.94	0.985	0.926	0.9845	0.94	0.989
	290	0.9465	0.9925	0.932	0.9865	0.946	0.9935

Figure **1**, it is evident that the Mean Squared Error (MSE) values for RSS are consistently lower compared to those obtained using Simple Random Sampling (SRS). Additionally, MSE values decrease as the failure rate increases, indicating improved estimation accuracy with more observed failures. When data are simulated from the Weibull Accelerated Failure Time (AFT) model, we observe that right censoring outperforms other censoring schemes based on RMSE values, producing the most efficient parameter estimates.

The coverage probabilities are summarized in Table **3**. Notably, left censoring performs competitively across all censoring schemes, achieving coverage probabilities that are consistently close to the desired 0.95 level, even when the failure rate is as low as 60%.

From Table **2**, we analyze the statistical power of the test as m/n increases for both sampling methods. While SRS exhibits higher power than RSS when $\beta_1 = 0$ and 0.05, the situation reverses when $\beta_1 = 1$. In this case, RSS outperforms SRS, achieving test power greater than 0.9 for both methods. These findings confirm that RSS provides more efficient parameter estimation than SRS, particularly when dealing with higher effect sizes and increasing failure rates.

Figures **1a**, **1b**, and **1c** illustrate the graphical representation of the Room Mean Square Error (RMSE*)* estimates of the Hazard Ratio (HR) under progressive type-II censoring from the Weibull AFT model. These figures are based on a sample of *n* = 300 with varying values of *m* and β_1.

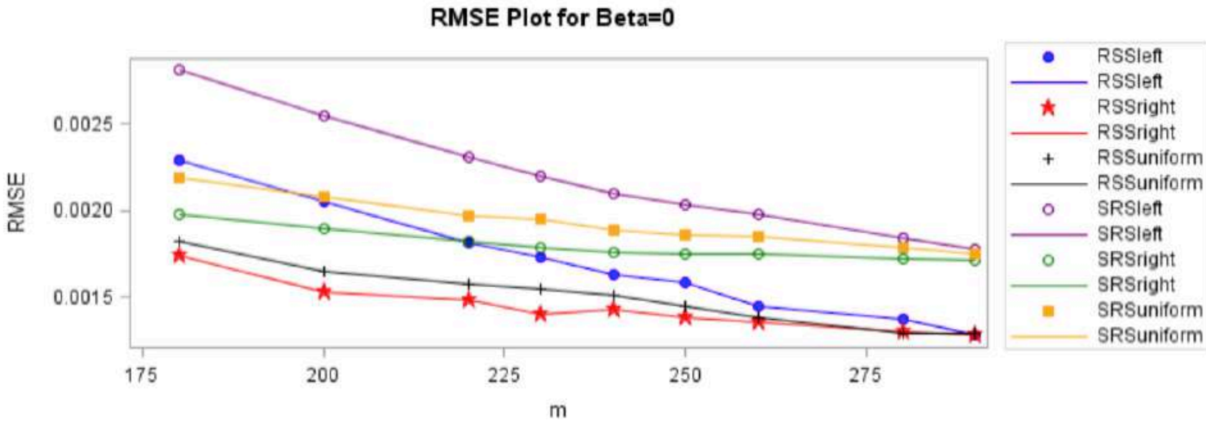
Figures **2a**, **2b**, and **2c** present the bias plots for the Hazard Ratio (HR) estimates under progressive Type-II censoring from the Weibull Accelerated Failure Time (AFT) model. These plots are based on a sample size of n = 300, with varying values of m and β_1.

5. ILLUSTRATION BASED ON ADVANCED LUNG CANCER DATA FROM NORTH CENTRAL CANCER TREATMENT GROUP

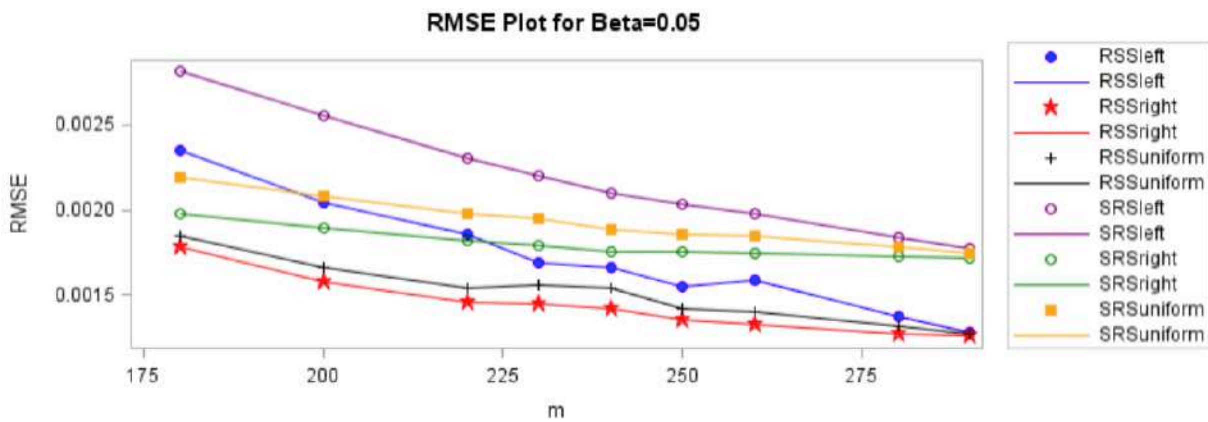
To demonstrate the effectiveness of our proposed method, we applied it to real-world survival data from the North Central Cancer Treatment Group (NCCTG). This data set consists of patients diagnosed with advanced lung cancer, including a performance score rating. Our goal is to obtain precise survival time estimates and better understand the relationship between survival time and patient characteristics.

5.1. Data Description and Preprocessing

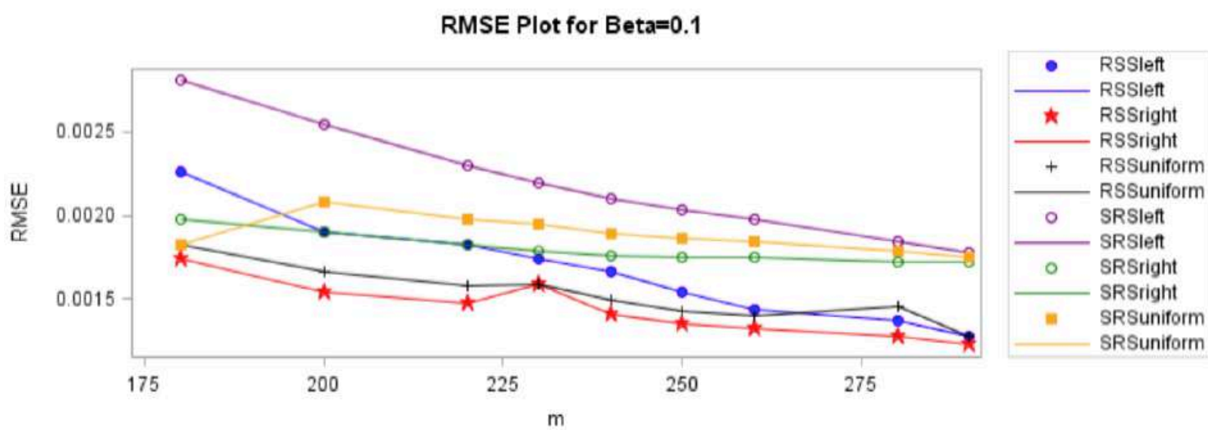
The dataset contains 228 individuals (138 men and 90 women) aged 35 to 82 years, with recorded survival times and associated covariates. The primary variables included in the analysis are:



a



b



c

Figure 1a: RMSE plot against m (events) when beta = 0 with Different Sampling Schemes.

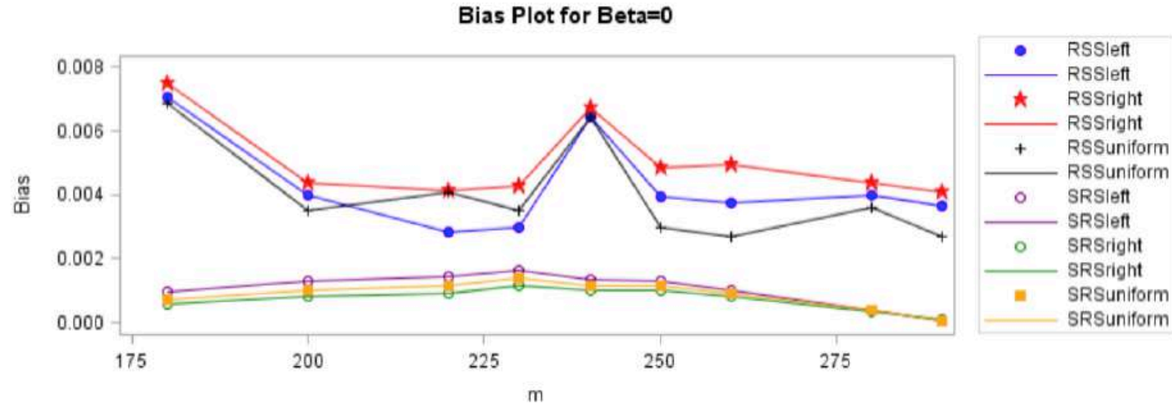
b: RMSE plot against m when beta = 0.05 with Different Sampling Schemes.

c: RMSE plot against m (events) when beta = 0.1 with Different Sampling Schemes.

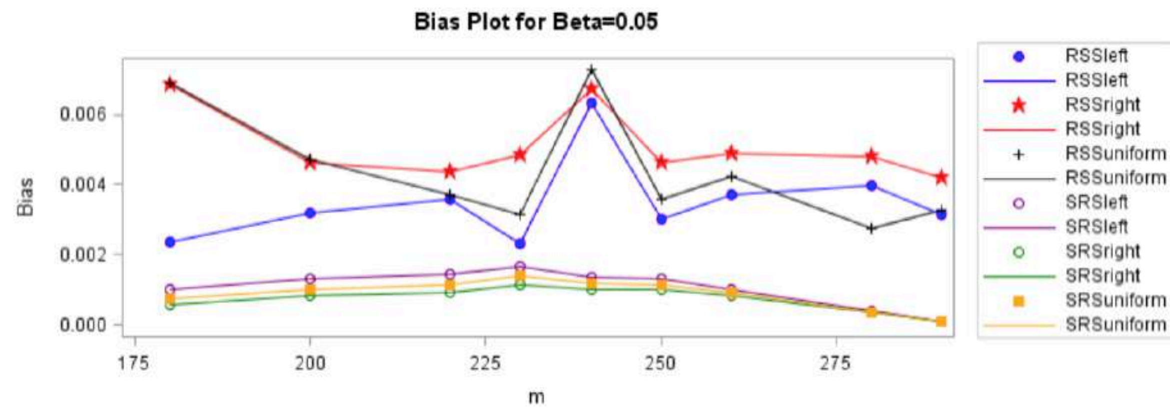
- Age: Patient's age in years
- Sex: Male or Female
- Survival time: Time (in days) from diagnosis to death or censoring
- Censoring status: Indicates whether survival time is observed (event) or censored.

The dataset is considered the population, and we drew a sample of 150 patients using Ranked Set Sampling (RSS) and Simple Random Sampling (SRS). The RSS selection was based on ranking individuals by age. Among the 150 sampled patients, 100 cases were observed as failures, while 50 were right-censored.

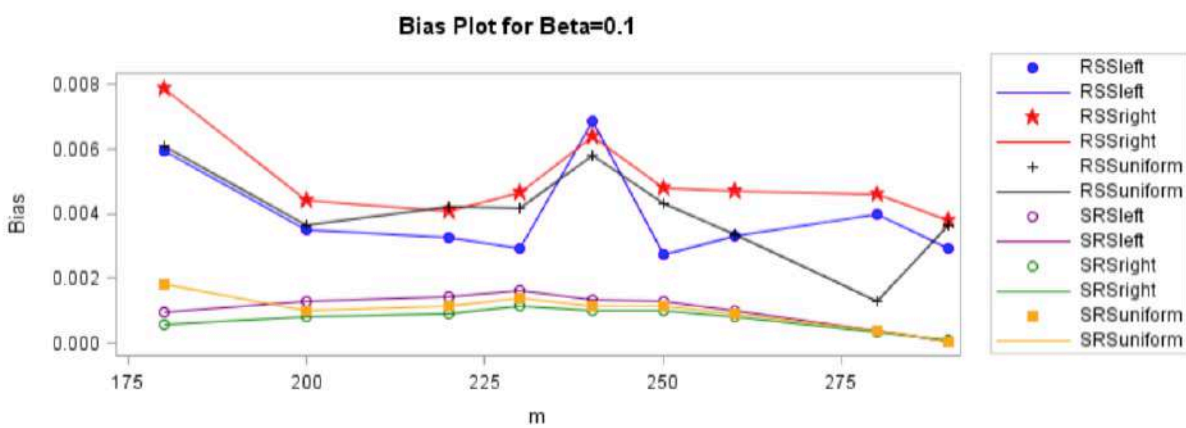
The dataset consists of patients diagnosed with advanced lung cancer and includes key characteristics



a



b



c

Figure 2a: Bias Plot for Beta = 0 with Different Sampling Schemes.

b: Bias Plot for Beta = 0.05 with Different Sampling Schemes.

c: Bias Plot for Beta = 0.1 with Different Sampling Schemes.

such as age, gender, survival time, and censoring status. The mean age of patients at diagnosis is 63.4 years, with a standard deviation of 9.2 years. Regarding gender distribution, 138 patients (60.5%) are male, while 90 patients (39.5%) are female. The median survival time from diagnosis to either death or censoring is 270 days. Additionally, 33.3% (50 cases) of the dataset are right-censored, indicating that survival times for these patients were not fully observed within the study period.

5.2. Model Selection and Weibull Distribution Validation

Before applying the Accelerated Failure Time (AFT) model, we tested whether the data follows a Weibull distribution, which is widely used in survival analysis.

5.2.1. Weibull Plot for Model Validation

To assess the Weibull fit, we utilized Weibull probability plotting, a graphical method to determine

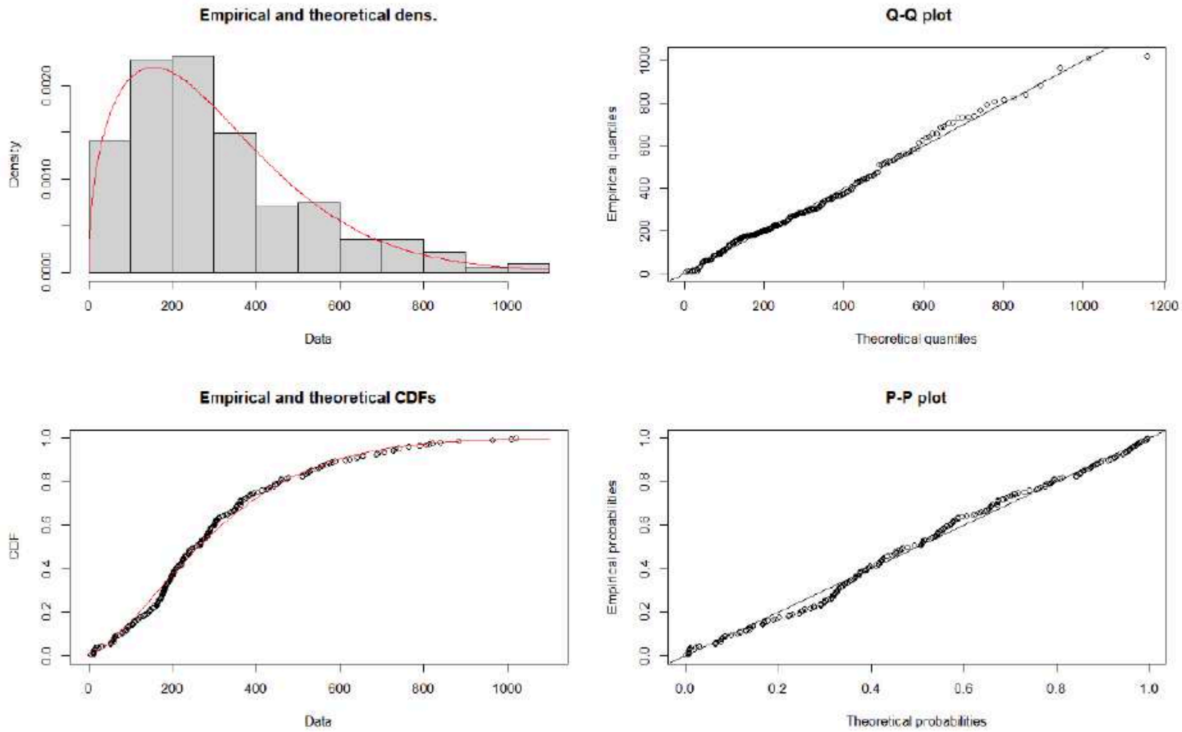


Figure 3*:* Diagnostics Fit for NCCTG Data.

whether survival times follow a two-parameter Weibull distribution. The empirical cumulative distribution function $\hat{F}$ is estimated using:

$$\hat{F} = \frac{i - 0.3}{n + 0.4} \quad (15)$$

where i is the rank of the data point, and *n* being the sample size.

For a complete samples, the Weibull plot consists of a Q-Q plot, where

$$y = log[-log(1 - F(t))]$$

is plotted against $log(T_{(i)})$ on the x-axis. A linear pattern in this plot indicates a good Weibull model fit.

For Type II censored samples, we only plot the first *r* observed failure times and exclude censored values while accounting for their presence in the total sample size.

Figure **3**. contains the Weibull plot (Q-Q plot) and some other plots for the data set to check which model best fits the data set.

5.2.2. Model Selection Using Information Criteria

To quantitatively confirm the best survival model, we computed the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for different survival distributions.

Table 4: Comparison of Model Fits using AIC and BIC

Model	AIC	BIC
Weibull	2311.702	2318.561
Log-Logistic	2325.861	2332.720
Log-normal	2342.538	2349.397
Exponential	2326.676	2330.106

The Weibull model has the lowest AIC and BIC, confirming that it provides the best fit for the survival data.

5.3. Applying the AFT Weibull Model: RSS vs. SRS

After confirming the Weibull model, we applied the AFT model under Progressive Type-II censoring, comparing results for both Ranked Set Sampling (RSS) and Simple Random Sampling (SRS).

5.3.1. Model Fit Statistics for RSS and SRS

5.3.1.1. Maximum Likelihood Parameter Estimates

The Ranked Set Sampling (RSS) method provides more precise estimates of survival time compared to Simple Random Sampling (SRS). Additionally, RSS demonstrates a better model fit, as indicated by its lower AIC and BIC values, and successfully identifies

Table 5: Model Fit Statistics - RSS vs. SRS

Methods	-2Log Likelihood	AIC	AICC	BIC
RSS	1397.839	1405.839	1406.114	1417.881
SRS	1410.463	1418.463	1418.739	1430.506

Table 6: Estimated Parameters for the AFT Weibull Model (RSS)

Analysis of Maximum Likelihood Parameter Estimates							
Parameter	df.	Estimate	Standard Error	95% Confidence Limits		Chi-Square	p-Value
Intercept	1	4.9931	0.5198	3.9743	6.0118	92.27	<.0001
age	1	0.0165	0.0083	0.0002	0.0327	3.92	0.0477
sex	1	0.1660	0.1255	-0.0799	0.4119	1.75	0.1858
Scale	1	0.6056	0.0452	0.5231	0.7010		

Table 7: Estimated Parameters for the AFT Weibull Model (SRS)

Analysis of Maximum Likelihood Parameter Estimates							
Parameter	df.	Estimate	Standard Error	95% Confidence Limits		Chi-Square	p-Value
Intercept	1	6.2987	0.5662	5.1889	7.4085	123.74	<.0001
age	1	-0.0067	0.0086	-0.0236	0.0102	0.60	0.4376
sex	1	0.1549	0.1515	-0.1421	0.4519	1.04	0.3067
Scale parameter	1	0.7247	0.0574	0.6205	0.8463		

age as a significant predictor of survival, whereas SRS fails to establish this relationship. These findings highlight the clinical utility of RSS, underscoring its potential to enhance efficiency and accuracy in survival analysis, making it a valuable tool in medical research and decision-making.

6. SENSITIVITY ANALYSIS

In this section, we perform a robust sensitivity analysis to assess the impact of varying sample sizes and censoring rates on the performance of two sampling techniques. The study employs three distinct sample sizes (100, 150, and 200) taken from the overall population of size 228 and three levels of censoring (20%, 40%, and 60%). Two sampling strategies are considered:

1. **Simple Random Sampling (SRS)** – A conventional approach where observations are randomly selected without ranking.
2. **Ranked Set Sampling (RSS)** – A method in which samples are pre-ranked based on an auxiliary variable (age) before selection.

To account for missing data, a proportion of observations (20%, 40%, and 60%) were randomly censored. The survival time was modeled as a function of age and sex using a Weibull Accelerated Failure Time (AFT) model. The performance of each sampling strategy was evaluated by computing MSE, Bias, AIC, and BIC across all experimental conditions.

6.1. Performance Comparison of RSS and SRS

A summary of the sensitivity analysis results is presented in Table **8**. The findings reveal that RSS consistently has lower MSE compared to SRS, indicating superior predictive accuracy. However, MSE increases with higher censoring rates, demonstrating the adverse impact of missing data on prediction reliability. Additionally, increasing the sample size reduces MSE, reinforcing the importance of adequate data collection for reliable estimation.

Missing data contributes to an increase in bias as censoring levels increase. While RSS displays slightly higher bias than SRS for smaller sample sizes, SRS exhibits significant deterioration in bias when censoring is more than 60%, highlighting its instability.

Table 8: Sensitivity Analysis of RSS and SRS

Method	Sample Size	Censoring Rate	Mean MSE	Mean Bias	Mean AIC	Mean BIC	Count
RSS	100	0.2	794.17	667.39	341.25	351.67	1000
RSS	100	0.4	7188.65	1516.12	251.89	262.31	1000
RSS	100	0.6	29096.81	6588.27	195.31	205.73	1000
RSS	150	0.2	537.19	611.90	545.89	557.93	1000
RSS	150	0.4	2278.42	1084.27	446.10	458.15	1000
RSS	150	0.6	50697.40	3597.77	307.02	319.06	1000
RSS	200	0.2	421.80	585.79	660.86	674.05	1000
RSS	200	0.4	1103.79	889.61	514.32	527.51	1000
RSS	200	0.6	15332.93	2571.48	313.34	326.54	1000
SRS	100	0.2	1342.78	204.97	41.65	42.88	1000
SRS	100	0.4	8772.95	1339.19	272.13	280.20	1000
SRS	100	0.6	44064.34	6726.42	1366.87	1407.39	1000
SRS	150	0.2	2677.43	408.71	83.05	85.51	1000
SRS	150	0.4	5630	859.42	174.64	179.81	1000
SRS	150	0.6	50425.36	7697.43	1564.19	1610.56	1000
SRS	200	0.2	433.75	66.21	13.45	13.85	1000
SRS	200	0.4	1110.88	169.45	540.90	554.09	1000
SRS	200	0.6	15401.33	2351.01	381.66	394.85	1000

AIC and BIC values consistently favor RSS across all sample sizes and censoring rates. Notably, SRS shows a substantial increase in AIC and BIC at the 60% censoring rate, further confirming its instability in datasets with high censoring.

The findings from this sensitivity analysis provide evidence that Ranked Set Sampling (RSS) consistently outperforms Simple Random Sampling (SRS) in survival analysis. RSS demonstrates superior predictive accuracy, evidenced by lower MSE values across all sample sizes and censoring rates. The method also exhibits more stable performance in the presence of censoring, as indicated by lower AIC and BIC values. Although bias increases with censoring, RSS remains more reliable, particularly at higher sample sizes.

7. DISCUSSION

The results demonstrate that the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values for the Ranked Set Sampling (RSS) method in the AFT Weibull model with Progressive Type-II (PII) censoring are lower than those obtained using Simple Random Sampling (SRS). This suggests that RSS provides a better model fit, making it a more efficient sampling strategy for estimating survival times.

Moreover, the analysis presented in Table **5** highlights a significant relationship between age and survival time when using RSS, whereas Table **6** indicates that in the SRS model, the intercept remains significant, but the effect of age is not statistically significant. This suggests that RSS offers a more precise estimation of the impact of age on survival compared to SRS.

These findings emphasize the advantages of Ranked Set Sampling (RSS) over Simple Random Sampling (SRS) in survival analysis. The improved efficiency of RSS allows for a more accurate estimation of covariate effects, particularly in cases where survival data is subject to censoring. Consequently, RSS proves to be a valuable alternative for enhancing parameter estimation in survival models, ultimately leading to better-informed clinical decisions.

8. PRACTICAL IMPLICATIONS AND LIMITATIONS

Ranked Set Sampling is not just a theoretical improvement over Simple Random Sampling; it has practical applications for real-world survival analysis and clinical research. One key advantage of RSS is its ability to optimize data collection in clinical trials when resources are limited. By ranking patients based on auxiliary variables such as performance scores or baseline health status, researchers can ensure that measured observations represent a broad spectrum of survival times. This sampling approach reduces the need for excessive recruitment. It allows for more targeted data collection, particularly useful in trials with long follow-up periods or high censoring rates.

Traditional survival models, such as the Accelerated Failure Time (AFT) model and the Cox Proportional Hazards model, can be adapted to account for the structure of RSS data. The sensitivity analyses conducted in this study demonstrate that RSS consistently yields lower Mean Squared Errors (MSE) and less bias compared to SRS. As a result, RSS can be a valuable tool for clinical trial design, post-market surveillance, and real-world data studies where efficient resource utilization is critical.

The effectiveness of Ranked Set Sampling (RSS) depends heavily on the choice of an appropriate auxiliary variable for ranking. If the selected auxiliary variable is poorly correlated with the variable of interest, the resulting sample may fail to be representative or informative, limiting the efficiency of the method. Additionally, because RSS involves a judgment-based ranking process, it is susceptible to subjectivity and potential bias, especially if rankings are inconsistent or influenced by individual raters' perspectives. Another practical limitation is that RSS requires additional information or auxiliary variables for ranking, which may not always be readily available. In cases where resources are limited, implementing RSS may be less feasible compared to simpler sampling methods.

9. CONCLUSION

Efficient sampling methods are crucial in statistics to reduce time and costs while maintaining reliable estimations. Progressive censoring is a valuable technique that improves estimator efficiency while minimizing resource use. In this study, we proposed using Ranked Set Sampling (RSS) for Progressive Type-II censored survival data to enhance estimation accuracy. The method was evaluated under the right, left, and uniform censoring schemes.

After confirming that the dataset follows a Weibull distribution, we applied the Accelerated Failure Time (AFT) model using maximum likelihood estimation for parameter estimation. The results showed that RSS consistently outperformed Simple Random Sampling (SRS), yielding better parameter estimates across different values of and under all censoring schemes. Additionally, coverage probabilities increased as the failure rate increased, demonstrating the robustness of the method. However, the simulation study revealed that as the failure rate increased, the test's statistical power tended to decrease. The sensitivity analysis highlights the limitations of SRS, especially at 60% censoring, where it shows a substantial increase in both MSE and bias, along with a significant rise in AIC and BIC, confirming model instability. These findings suggest that RSS should be the preferred sampling method in biostatistical survival analysis studies, mainly when auxiliary ranking information is available.

These findings highlight the advantages of using RSS in survival analysis, particularly for Progressive Type-II censoring scenarios. The method enhances estimation accuracy and efficiency, making it a promising alternative for future medical and reliability studies applications.

REFERENCES

[1] Balakrishnan N, Cramer E. The art of progressive censoring : applications to reliability and quality 2014; 645.

[2] Helu A, Samawi H. On Marginal Distributions under Progressive Type II Censoring: Similarity/Dissimilarity Properties. Open J Stat 2017; 07: 633-44. https://doi.org/10.4236/OJS.2017.74044

[3] Biradar BS, Santosha CD, Wolfe DA. Ranked Set Sampling An Approach to More Efficient Data Collection. Statistical Science 2004; 19: 636-643. https://doi.org/10.1214/088342304000000369

[4] Balakrishnan N, Aggarwala R. Progressive Censoring. Progressive Censoring 2000. https://doi.org/10.1007/978-1-4612-1334-5

[5] Helu A, Samawi H, Alslman M. On Accelerated Failure Time Models Performance Under Progressive Type-II Censoring. Statistics, Optimization & Information Computing 2022; 10: 815-28. https://doi.org/10.19139/soic-2310-5070-1446

[6] Estimating the Population Mean Using Extreme Ranked Set Sampling n.d. https://www.researchgate.net/publication/242939607_Estimating_the_Population_Mean_Using_Extreme_Ranked_Set_Sampling (accessed January 12, 2023).

[7] Samawi H, Al-Saleh M, Saidy O. (PDF) On the matched pairs sign test using bivariate ranked set sampling: an application to environmental issues 2008. https://www.researchgate.net/publication/228675925_On_the_matched_pairs_sign_test_using_bivariate_ranked_set_sampling_an_application_to_environmental_issues (accessed February 12, 2023).

[8] Samawi H, Pararai M. On the optimality of bivariate ranked set sample design for the matched pairs sign test on JSTOR 2010:24–7. https://www.jstor.org/stable/43601141 (accessed February 9, 2023).

[9] Samawi HM, Al-Saleh MF. On regression analysis using bivariate ranked set samples. Metron 2002; 60: 29-48.

[10] Samawi H, Al-Saleh M. On regression analysis using bivariate ranked set samples - Georgia Southern University 2002:29–48. https://scholars.georgiasouthern.edu/en/publications/on-regression-analysis-using-bivariate-ranked-set-samples-4 (accessed June 2, 2023).

[11] Samawi HM, Ahmed MS, Abu-Dayyeh W. Estimating the population mean using extreme ranked set sampling. Biometrical Journal 1996; 38: 577-86. https://doi.org/10.1002/BIMJ.4710380506

[12] El-Sayed MA, Riad FH, Elsafty MA, Estaitia YA. Algorithms of Confidence Intervals of WG Distribution Based on Progressive Type-II Censoring Samples. Journal of Computer and Communications 2017; 05: 101-16. https://doi.org/10.4236/JCC.2017.57011

[13] Elhag AA, Ibrahim OIO, El-Sayed MA, Abd-Elmougod GA. Estimations of Weibull-Geometric Distribution under Progressive Type II Censoring Samples. Open J Stat 2015; 05: 721-9. https://doi.org/10.4236/OJS.2015.57072

[14] Mahmoud MAW, Soliman AA, Ellah AHA, El-Sagheer RM. Estimation of Generalized Pareto under an Adaptive Type-II Progressive Censoring. Intell Inf Manag 2013; 05: 73-83. https://doi.org/10.4236/IIM.2013.53008

[15] Loprinzi CL, Laurie JA, Wieand HS, Krook JE, Novotny PJ, Kugler JW, *et al*. Prospective evaluation of prognostic variables from patient-completed questionnaires. North Central Cancer Treatment Group. J Clin Oncol 1994; 12: 601-7. https://doi.org/10.1200/JCO.1994.12.3.601

[16] Balakrishnan N, Cramer E. The art of progressive censoring: Applications to reliability and quality. The Art of Progressive Censoring: Applications to Reliability and Quality 2014; 1-645. https://doi.org/10.1007/978-0-8176-4807-7

[17] Balakrishnan N, Cramer E. Statistics for Industry and Technology The Art of Progressive Censoring Applications to Reliability and Quality 2014.

Received on 02-01-2025 Accepted on 10-02-2025 Published on 02-03-2025

https://doi.org/10.6000/1929-6029.2025.14.09