Further Increasing Fisher's Information for Parameters of Association in Accelerated Failure Time Models via Double Extreme Ranks

Amal Helu
Deaprtment of Mathematics
The University of Jordan
Amman 11942, Jordan
a.helu@ju.edu.jo

Hani M. Samawi
Department of Biostatistics
Jiann-Ping Hsu College of Public Health,
Georgia Southern University, USA
hsamawi@georgiasouthern.edu

Haresh Rochani
Department of Biostatistics
Jiann-Ping Hsu College of Public Health,
Georgia Southern University, USA
hrhchani@georgiasouthern.edu

Abstract

Double Extreme Ranked Set Sampling (DERSS) was first introduced by Samawi (2002) as a modification to the well-known Ranked Set Sampling (RSS) and Extreme Ranked Set Sampling (ERSS). In this article, we provide a modification to DERSS scheme with ranking based on an easy-to-evaluate baseline auxiliary variable known to be associated with survival time. We show that using the modified DERSS improves the performance of the Accelerated failure time (AFT) survival model and provides a more efficient estimator of the hazard ratio than that based on their counter parts simple random sample (SRS), RSS and ERSS. Our theoretical and simulation studies show the superiority of using the modified DERSS for AFT survival models compared with using SRS, RSS and ERSS. A numerical example based on Worcester Heart Attack Study is presented to illustrate the implementation of the DERSS.

Keywords: Accelerated Failure Time Model, Hazards Ratio, Double Extreme Ranked Set Sampling, Extreme Ranked Set Sampling, Survival analysis.

Introduction

Survival analysis can be used to evaluate the effects of covariates on the time until a subject experiences the event of the study. Some examples of events include: death, appearance of a tumor, development of some disease, recurrence of a disease, conception, or cessation of smoking. The most popular methods to analyze the effect of covariates on survival time are regression models. One class of regression models for survival data is the Cox proportional hazards (PH) models (Cox, 1972). The other class of popular modeling for survival data is the parametric AFT models (Collett, 2003). While Cox proportional hazards models relate the hazard function to covariates, the AFT models specify a direct relationship between the failure time and covariates.

AFT models are especially useful in crucial industrial applications in which failure time is accelerated, for example, by thermal high-voltage or other factors. In addition, the AFT models are occasionally used to analyze survivorship data of elderly people. Moreover, AFT modeling is based upon the survival curve but not on the hazard function (Swindell, 2009). Furthermore, the applications and development of the proportional hazards model rely on the assumption of independent responses from the monitored units that are subjected to failure. AFT models share similar assumptions (Chapman et al., 1992).

One of the common sampling designs for statistical inference is simple random sampling (SRS) (Scheaffer et al., 2006). In many studies a large sample size, if SRS used, to have a representative sample to the population is needed to validate the study results. An alternative cost effective sampling scheme to SRS is ranked set sampling (RSS). RSS was first proposed by McIntyre (1952) and used in agriculture studies. RSS is more structural sampling scheme than SRS and provides more representative sample to the study population due to its natural stratification behavior based on the order statistics. Hence RSS needs less quantified sampling units than SRS to achieve the same accuracy for many statistical procedures.

The balanced RSS procedure starts with randomly drawing m^2 subjects from the target population then these subjects are randomly divided into *m* sets of *m* subjects each. Within each set, the subjects are ranked usually visually or by the mean of an available concomitant (auxiliary) variable (say *Z*), such as age, BMI and weight, related to the variable of interest (say *Y*). From the first set of *m* subjects, the subject with the lowest rank with respect to (*Y* or *Z*) is chosen for actual measurement of the variable of interest (*Y*). From the second set of *m* subjects, the subject ranked second lowest is measured. The process is continued until the subject ranked highest is measured from the m^{th} set of *m* subjects. The set size *m* needs to be small no more than five due to the fact that larger set size may cause error in the ranking process. If larger sample size is needed the procedure above can be repeated *r* time to have a sample size $n=r.m$. We call *r* the cycle size.

Ranked set sampling is prone to ranking error if the chosen set size *m* is large. To overcome this problem, several variations of RSS have been proposed by researchers. Samawi et al. (1996), investigated the performance of ERSS in estimating the population mean. Another modification of RSS, namely DERSS, was introduced by Samawi (2002) for mean and regression estimators. A modified version of ERSS can be implemented using then following steps:

1) Randomly drawing *m* independent sets each contains *m* sample units;

2) We assume that the maximum or the minimum sample unit within each set with respect to the value of an auxiliary variable Z, which is associated with survival time, can be identified with no or little cost. Then order the sampling units in each set with respect to available baseline auxiliary variable (Z).

3) Measure the maximum (or minimum) ordered unit from each set. This sampling scheme is known as $ERSS_{max}$ (or $ERSS_{min}$) of size *m* (Samawi et al., 2018). Furthermore, to draw a modified DERSS, use step 1 to 3, which describing the implementation of the modified ERSS above to, without actual measurements, select *m* $ERSS_{max}$ (or $ERSS_{min}$). Then measure the maximum (or minimum) from each. The propose sampling scheme is called $DERSS_{max}$ (or $DERSS_{min}$). If large sample size is needed the whole cycle can be repeated *r* times to have a sample of size $n=r.m$. In fact, the modified ERSS and DERSS still inheriting the stratification behavior of RSS. However, focusing on targeting the upper or the lower part of the population, which increase the number of the event's occurrences, improves the performance of the AFT models.

Most of RSS and its variations sampling schemes are implemented in agriculture and environmental studies. However, Samawi and Al-Sagheer (2001) were first to apply

RSS in a study involving human subjects. They described the data collection in the study that involved the analysis of the level of bilirubin in the blood of the jaundice premature babies. Samawi and Al-Sagheer (2001) suggested that an expert physician could do ranking on the level of bilirubin in the blood visually by observing: i) Color of the face ii) Color of the chest iii) Color of lower part of the body iv) Color of terminal parts of the whole body. As the level of bilirubin in the blood increases, the yellowish discoloration goes from i). to iv.). Also, see Jabrah et al. (2017) for another application of RSS in a study involving human subjects. They applied the RSS sampling design to select college students for the analysis of a psychological intervention to buttress resilience study. On the other hand ERSS became a useful sampling scheme in some medical fields. For example, ERSS sampling scheme recently applied to genetics for quantitative trait loci (QTL) mapping (Chen, 2007). Chen (2007) explained that in case of the frequency of the Q allele, in the general population is small, one of the alternatives approaches adopted to draw SRS for detecting QTL using population data is to truncate the population at a certain quantile of the distribution of response variable (Y) and take a random sample from the truncated portion and a random sample from the whole population. The two samples drawn are genotyped and compared on the number of Q-alleles. Then if a significant difference exists, the candidate QTL is claimed as a true QTL (Chen 2007). However, this approach needs a large number of individuals have to be screened before a sample can be taken from the truncated portion and hence it is not practical. Alternatively, the ERSS is used as follows: Individuals are taken in sets and the individuals within each set are ranked according to their trait values. The one with the largest trait value is put into an upper sample and the one with the smallest trait value is put into a lower sample. Then the two samples obtained this way are then genotyped and compared. Also, ERSS approach has been applied for linkage disequilibrium mapping of QTL recently by Chen et al. (2005). The ERSS has been applied to a sib-pair regression model where extremely concordant and/or discordant sib-pairs are selected by the ERSS (see Zheng et al. 2006). As indicated by Chen (2007), the ERSS approach can be applied also to many other genetic problems such as the transmission disequilibrium test (TDT) and the gamete competition model (Sinsheimer et al. 2000).

For improving the inference of regression models, RSS has recently gained significant consideration as an efficient sampling design. For example, Samawi and Ababneh (2001), implemented RSS based on ranking on the covariate (X) to investigate its effect on regression analysis. Samawi and Abu-Dayyeh (2002) further extended this work by assuming the regressors to be random. Rochani et al. (2018) demonstrated that the efficiency of multivariate regression estimator can be improved by using RSS. The literature on this topic is extensive in the last 50 years, for example see (Al-Saleh and Samawi, 2000; Al-Saleh and Zheng, 2003; Samawi and Al-Saleh, 2002, Samawi et al., 2009, Samawi et al., 2018.) In addition, Samawi et al. (2018) improved the performance of AFT survival model by using a modified ERSS, namely $ERSS_{min}$ or $ERSS_{max}$.

The first aim of this paper is to introduce $DERSS_{min}$ ($DERSS_{max}$) sampling scheme, which is an extension to the modified $ERSS_{min}$ ($ERSS_{max}$) scheme. The derivation of the sampling distribution using $DERSS_{min}$ ($DERSS_{max}$) is provided in the next section.

The second aim in this paper, is to show, theoretically and by simulation, that using the modified DERSS improves the performance of AFT survival models and provides more efficient estimators of the hazard ratios compares with their counter parts, simple random sample (SRS), RSS and ERSS. The remainder of this paper is organized as follows: In Section 2 we introduce sample notations and some basic results about the modified

DERSS. AFT regression model and its properties using the modified DERSS will be discussed in section 3. In Section 4, we provide a simulation study to compare the performance of DERSS, ERSS and SRS for all the AFT models. In Section 5, we illustrate the method using Worcester Heart Attack Study. Final remarks are given in Section 6.

Preliminaries

Let T be a random variable for time to an event. The distribution of T is usually described or characterized by three functions, namely: the survival function, denoted by $S(t)$; the hazard rate function or risk function, denoted by $h(t)$; and the probability density (or probability mass) function, denoted by $f(t)$. The unique feature about survival data is censoring. Censored data arise when exact time to event for a subject is unknown. There are several censoring methods available to researchers, for example; Type I censoring in which the test ceases at a prefixed time, or Type II censoring that allows the experiment to be terminated at a predetermined number of failures. In this paper, we will only focus on Type I right censoring.

For the i^{th} individual, the lifetime observation can be described by (t_i, γ_i), where t_i is the survival time, which can be defined as $t_i = \min(T_i, C_i)$. Hence, T_i is the true survival time and C_i is the censoring time. As in Samawi et al. (2018), define γ_i as the indicator variable for survival status as follows:

$$\gamma_i = \begin{cases} 1 \text{ if } T = t_i \text{ or } T \le C_i & \text{(uncensored)} \\ 0 \text{ if } T > t_i \text{ or } T > C_i & \text{(right censored).} \end{cases} \quad (1)$$

Therefore, given lifetimes $t_1, t_2, ..., t_n$ for a SRS of n individuals, the likelihood function for the sample is given by

$$l(\boldsymbol{\beta}) = \prod_{i=1}^{n} f(t_{\mathbf{i}})^{\gamma_i} S(t_i)^{1-\gamma_i}, \quad (2)$$

where $\boldsymbol{\beta}$ is the vector of parameters to be estimated in the presence of right censoring.

Sample Notation and Some Basic Results of the Modified DERSS

For the k-th cycle, let $Z^1_{11k}, Z^1_{12k}, ..., Z^1_{1mk}, Z^1_{21k}, Z^1_{22k}, ..., Z^1_{2mk}; ...; Z^1_{m1k}, Z^1_{m2k}, ..., Z^1_{mmk};$ $Z^m_{11k}, Z^m_{12k}, ..., Z^m_{1mk}, Z^m_{21k}, Z^m_{22k}, ..., Z^m_{2mk}; ...; Z^m_{m1k}, Z^m_{m2k}, ..., Z^m_{mmk};$ $k=1, 2, ..., r$, be the m independent sets each with sample size m^2. Note that Z^l_{ijk} is the j-th sample unit in the i-th row (sample) of the k-th set. Assume that each element Z^l_{ijk} in the sample has a p.d.f. $f_Z(z)$ and a distribution function $F_Z(z)$ (absolutely continuous). Selecting DERSS$_{\min}$ after ranking the sample units within each sample in each set (visually or by any non costly way) we obtain:

$$\begin{bmatrix} Z^1_{1(1)k}, & Z^1_{1(2)k}, & ..., & Z^1_{1(m)k} \\ Z^1_{2(1)k}, & Z^1_{2(2)k}, & ..., & Z^1_{2(m)k} \\ ... & ... & ... & ... \\ Z^1_{m(1)k}, & Z^1_{m(2)k}, & ..., & Z^1_{m(m)k} \end{bmatrix}, \ldots, \begin{bmatrix} Z^m_{1(1)k}, & Z^m_{1(2)k}, & ..., & Z^m_{1(m)k} \\ Z^m_{2(1)k}, & Z^m_{2(2)k}, & ..., & Z^m_{2(m)k} \\ ... & ... & ... & ... \\ Z^m_{m(1)k}, & Z^m_{m(2)k}, & ..., & Z^m_{m(m)k} \end{bmatrix},$$

$k=1, 2, ..., r$. Thus the first stage will yield m ERSS$_{\min}$ samples:

A_{1k} $=\{ Z^1_{1(1)k}, Z^1_{2(1)k}, ..., Z^1_{m\text{-}1(1)k}, Z^1_{m(1)k} \}$,

A_{2k} $=\{ Z^2_{1(1)k}, Z^2_{2(1)k}, ..., Z^2_{m\text{-}1(1)k}, Z^2_{m(1)k} \}$, ...,

A_{mk} $=\{ Z^m_{1(1)k}, Z^m_{2(1)k}, ..., Z^m_{m\text{-}1(1)k}, Z^m_{m(1)k} \}$. Now let $V_{1(1)k}=\min(A_{1k})$, $V_{2(1)k}=\min(A_{2k})$, ..., $V_{m(1)k}=\min(A_{mk})$ Then $V_{1(1)k}, V_{2(1)k}, ..., V_{m(1)k}$ *k=1, 2, ..., r*, denotes DERSS$_{min}$. The DERSS$_{max}$
is similar to DERSS$_{min}$ but the maximums are selected instead of the minimums. It is easy to show that the p.d.f of the smallest and the largest order statistics of an i.i.d sample of size *m* with p.d.f $f_Z(z)$ are respectively given by $f_{Z(1)}(z)=m(1-F_Z(z))^{m-1} f_Z(z)$ and $f_{Z(m)}(z)=m(F_Z(z))^{m-1} f_Z(z)$. Also, let $V_{i(1)k}$ have p.d.f $g_{(1)}(z)$ and c.d.f $G_{(1)}(z)$ and $V_{i(m)k}$ have p.d.f $g_{(\mathrm{m})}(z)$ and c.d.f $G_{(\mathrm{m})}(z)$ where *i=1,2,...,m and k=1, 2, ..., r*. Clearly, $V_{1(1)k}, V_{2(1)k}, ..., V_{m(1)k}$, *k=1, 2, ..., r*, are independent and identically distributed. Using the above description of DERSS, we have the following lemmas:

Lemma (1): Under the above assumption,

(1) $G_{V_{(1)}}(x)=1-[1-F_{Z_{(1)}}(z)]^m=1-[1-F_Z(z)]^{m^2}$,

(2) $\mathrm{g}_{V_{(1)}}=mf_{Z_{(1)}}(z)[1-F_{Z_{(1)}}(z)]^{m-1}=\mathrm{m}^2 f_Z(z)\,[1-F_Z(z)]^{m^2-1}$

(3) $G_{V_{(m)}}(z)=[F_Z(z)]^{m^2}$

(4) $\mathrm{g}_{V_{(\mathrm{m})}}(z)=\mathrm{m}^2 f_Z(z)\,[F_Z(z)]^{m^2-1}$

Proof: (1) and (3) can be shown directly from the definition of the c.d.f of the random variables
$V_{i(1)k}$ and $V_{i(m)k}$ respectively. Also, (2) and (4) can be shown by taking the first derivative, with respect to z, of (1) and (3) respectively.

Moreover, in this procedure only the maximum (or minimum) of sets of a fixed size is identified for quantification. Therefore, even for large m, the modified DERSS can be easily implemented. We can allow for a larger set size m by ranking the sampled units based on an auxiliary variable (Z) that is highly correlated with the variable of interest.

AFT Model Using DERSS$_{min}$

In this section we derive the AFT models properties under DERSS$_{min}$ and show how using DERSS$_{min}$ improves the performance of the AFT Models. We assume that the relation between the survival time (*T*) and the ranked auxiliary variable (*Z*), which is assumed to be easy to be ranked, is positive. However, when the relation between the survival time and the auxiliary variable is negative we suggest to use DERSS$_{max.}$ The derivation of using DERSS$_{max}$ is similar and will not be provided in this paper.

For $k=1,2,...,r$, let $V_{1(1)k}, V_{2(1)k}, ..., V_{m(1)k}$ be the measurements of DERSS$_{min}$ of size $n=r.m$, obtained based on ranking the auxiliary variable (*Z*). We assume that the judgment on ranking of the auxiliary variable is perfect. Therefore, for the kth cycle, $V_{(1)k}$ is defined as the minimum order statistic of an ERRS$_{min}$ of size *m*. Now as in Samawi et al. (2018), define the vector $\mathbf{w}_{i[1]k}=(x_{i[1]k}...x_{i[1]kp}, v_{i(1)k})'$, $i=1,2,...,m; k=1,2,...,r$. Clearly, the vector $\mathbf{w}_{i[1]k}$ represents the observations on the *p+1* explanatory variables plus one auxiliary variable. Note that as indicated by Samawi et al. (2018), the notation (.) used for perfect ranking while the notation [.] used for imperfect ranking.

Similar to Liu (2012) and Samawi et al. (2018), the log-linear form of the AFT model, with respect to $\log T_{i[1]k}$, is given

$$log\, T_{i[1]k} = \beta_0 + x_{i[1]k1}\beta_1 + \ldots + x_{i[1]kp}\beta_p + v_{i(1)k}\beta_{p+1} + \sigma\varepsilon_{ij}, i = 1,2,\ldots,m; k = 1,2,\ldots,r. \tag{3}$$

Note that, σ is a scale parameter and ε_{ij} is a random error term of the model which is assumed to have a specific distribution. Adopting the same notation as in Samawi et al. (2018), the survival function at time $T_{i[1]k}$ is as follows:

$$\begin{aligned} S_{ik}(t) &= P(log\, T_{i[1]k} \geq log\, t_{i[1]k}) \\ &= P\left(\varepsilon_{ik} \geq \frac{log\, t_{i[1]k} - \{\beta_0 + x_{i[1]k1}\beta_1 + \ldots + x_{i[1]kp}\beta_p + v_{i(1)k}\beta_{p+1}\}}{\sigma}\right) \\ &= P\left(\varepsilon_{ik} \geq \frac{log\, t_{i[1]k} - \beta_0 - \boldsymbol{w}'_{i[1]k}\boldsymbol{\beta}}{\sigma}\right). \end{aligned} \tag{4}$$

Assuming $\mathbf{w}_{i[1]k} = (x_{i[1]k1},\ldots,x_{i[1]kp},v_{i(1)k})',\ i = 1,2,\ldots,m; \mathrm{k} = 1,2,\ldots,r$ fixed, the survival function of the $\log T_{i[1]k}$ is given by

$$\begin{aligned} S_{ij}(t) &= P(\log T_{i[1]k} \geq \log t_{i[1]k}) \\ &= S_0\left(\frac{\log t_{i[1]k} - \beta_0 - \mathbf{w}'_{i[1]k}\boldsymbol{\beta}}{\sigma}\right), -\infty < t_{i[1]k} < \infty, \\ &= P\left(\varepsilon_{ik} \geq \frac{\beta_0 + x_{i[1]k1}\beta_1 + \ldots + x_{i[1]kp}\beta_p + v_{i(1)k}\beta_{p+1}}{\sigma}\right) \\ &= P\left(\varepsilon_{ik} \geq \frac{\log t_{i[1]k} - \beta_0 - \mathbf{w}'_{i[1]k}\boldsymbol{\beta}}{\sigma}\right) \end{aligned} \tag{5}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_{2,},\ldots,\beta_{p+1})'$.

Furthermore, the hazard function for $T_{i[1]k}$ under DERSS$_{\min}$ scheme is

$$h_{ik}(t \mid \mathbf{w}_{i[1]k}, \boldsymbol{\beta}) = \frac{1}{t_{i[1]k}\sigma} h_0\left(\frac{\log t_{i[1]k} - \beta_0 - \mathbf{w}'_{i[1]k}\boldsymbol{\beta}}{\sigma}\right),\ i = 1,2,\ldots,m; k = 1,2,\ldots,r, \tag{6}$$

where $h_0(t)$ is the baseline hazard function at survival time t. As in Samawi et al. (2018), the covariates { $X_{i[1]k1},\ldots,X_{i[1]kp},V_{i(1)k}$ } are assumed to have a multiplicative effect on the hazard function. Therefore, the predicted value of hazard function, given

{ $x_{i[1]k1},...,x_{i[1]kp},v_{i(1)k}$ }, which denoted by $\hat{h}(t_{i[1]k} \mid x_{i[1]k1},...,x_{i[1]kp},v_{i(1)k})$, can take values in the range $(0,\infty)$.

The Likelihood Function

Given n independent observations $(t_{i[1]k},\gamma_{ik},\mathbf{w}_{i[1]k} : i=1,2,...,m, K=1,2,...r)$, the likelihood function can be written as (see Liu, 2012):

$$l(\boldsymbol{\beta}) = \prod_{i=1}^{m}\prod_{k=1}^{r} f(t_{i[1]k})^{\gamma_{ik}} S(t_{i[1]k})^{1-\gamma_{ik}}. \qquad (7)$$

Then the log likelihood function has the following form,

$$L(\boldsymbol{\beta}) = \log[l(\boldsymbol{\beta})] \propto \sum_{i=1}^{m}\sum_{k=1}^{r}\left\{\left[\gamma_{ik}\log f(t_{i[1]k})\right] + (1-\gamma_{ik})\log S(t_{i[1]k})\right\}. \qquad (8)$$

An iterative procedure, such as, Newton-Raphson methods, to obtain MLE estimates of the p+1 unknown parameters $\boldsymbol{\beta} = (\beta_1,\beta_{2,},...,\beta_{p+1})'$ is needed.

To estimate the unknown parameters based on the $L(\boldsymbol{\beta})$, we need to differentiate the $\{L(\boldsymbol{\beta})\}$ with respect to β_j . The first derivative of the j^{th} element of the vector of parameters, $\boldsymbol{\beta}$ is:

$$\breve{S}(\beta_j) = \frac{dL(\boldsymbol{\beta})}{d\beta_j} \qquad (9)$$

where $\breve{S}(\beta_j)$ is called the score. The vector $\boldsymbol{\beta}$ can be estimated by solving the equation

$$\breve{S}(\boldsymbol{\beta}) = 0, \qquad (10)$$

where $\breve{S}(\boldsymbol{\beta}) = (\partial L/\partial\beta_0,....,\partial L/\partial\beta_{p+1})'$.

For large samples, $\hat{\boldsymbol{\beta}}$ is a unique solution of $\breve{S}(\boldsymbol{\beta}) = 0,$ and $\hat{\boldsymbol{\beta}}$ is consistent for $\boldsymbol{\beta}$ and has a multivariate normal distribution, which implies

$$\hat{\boldsymbol{\beta}} \sim MN(\boldsymbol{\beta},\widehat{V}(\boldsymbol{\beta})). \qquad (11)$$

We need to solve these equations iteratively using a numerical technique such as the Newton-Raphson method (See, Agresti, 2002). Under the same regularity conditions, used in SRS, the resulting solution of Newton-Raphson approximation converges to the maximum likelihood estimates of $\hat{\boldsymbol{\beta}}$ (Liu, 2012). That is because the log likelihood function in (8) is strictly concave in $\boldsymbol{\beta}$, then the maximum likelihood estimates (MLE) of βs do exist and they are unique except in certain boundary cases (Agresti, 2002). Lehmann and Casella (1998), provided six regularity conditions for maximum likelihood estimators to be consistent. Many authors, including Rashid and Shifa (2009), showed that all those regularity conditions are satisfied for AFT type of models. Also they showed that under those regularity conditions, the MLE of βs are consistent estimators and asymptotically normally distributed. Similarly, under the same regularity conditions, the resulting solution of Newton-Raphson approximation converges to the maximum likelihood estimates of $\hat{\boldsymbol{\beta}}$ (Albert and Anderson, 1984) and hence they are consistent estimators.

Moreover, the asymptotic normal distribution of $\hat{\boldsymbol{\beta}}$ provides the basis for hypothesis testing on $\boldsymbol{\beta}$. The variance covariance matrix estimator $\widehat{V}(\boldsymbol{\beta})$ of $\hat{\boldsymbol{\beta}}$ is based on the *(P+2) by (P+2)* observed information matrix $\mathbf{I}(\hat{\boldsymbol{\beta}})$. In this case, $\widehat{V}(\boldsymbol{\beta}) = \mathbf{I}(\hat{\boldsymbol{\beta}})^{-1}$. Now, to obtain $\mathbf{I}(\boldsymbol{\beta})$ we need the second derivative of the log-likelihood function and then

$$\mathbf{I}(\boldsymbol{\beta})_{DERSS_{\min}} = -E_{V_{(1)}} E_{T_{[1]}|V_{(1)}} \left(\frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_j \partial \beta_{j'}} \right)_{(P+2)x(P+2)}.$$

AFT Models

AFT- Exponential Regression Model

For the exponential AFT model the hazard function under this model is constant over time. Therefore, the hazard function can be written as

$$h_{ik}(t \mid \mathbf{w}_{i[1]k}, \boldsymbol{\beta}) = \lambda \exp\left(-\mathbf{w}'_{i[1]k}\boldsymbol{\beta}\right),\ i = 1,2,...,m; k = 1,2,...,r\ . \tag{12}$$

If we view $\log \lambda$ as a coefficient and place it into the regression coefficients vector $\boldsymbol{\beta}$ then (12) can be simplify to

$$h_{ik}(t \mid \mathbf{w}_{i[1]k}, \boldsymbol{\beta}) = \exp\left(-\mathbf{w}'_{i[1]k}\boldsymbol{\beta}\right),\ i = 1,2,...,m; k = 1,2,...,r\ . \tag{13}$$

Thus, the survival function given the exponential distribution of the event time T is given by

$$S_{ik}(t) = \exp\left[-\exp\left(\log t_{i[1]k} - \mathbf{w}'_{i[1]k}\boldsymbol{\beta}\right)\right], -\infty < \log t_{i[1]k} < \infty. \tag{14}$$

The density function is given by

$$f(t_{ik}) = \exp\left[\left(\log t_{i[1]k} - \mathbf{w}'_{i[1]k}\boldsymbol{\beta}\right) - \exp\left(\log t_{i[1]k} - \mathbf{w}'_{i[1]k}\boldsymbol{\beta}\right)\right], -\infty < \log t_{i[1]k} < \infty \tag{15}$$

Then the log likelihood function is given by

$$L(\boldsymbol{\beta}) = \log[l(\boldsymbol{\beta})] \propto \sum_{i=1}^{m}\sum_{k=1}^{r}\left[\gamma_{ik}(\log t_{i[1]k} - \mathbf{w}'_{i[1]k}\boldsymbol{\beta}) - t_{i[1]k}\exp\left(-\mathbf{w}'_{i[1]k}\boldsymbol{\beta}\right)\right]. \tag{16}$$

For the j^{th} covariate, the MLE of β_j is given by solving

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_j} = \sum_{i=1}^{m}\sum_{k=1}^{r}\left[-x_{ji[1]k}\{\gamma_{ik} - t_{i[1]k}\exp\left(-\mathbf{w}'_{i[1]k}\hat{\boldsymbol{\beta}}\right)\}\right] = 0.$$

However, the second partial derivative of the log-likelihood function, used to obtain the Fisher information matrix

$$-\left[\left(\frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_j \partial \beta_{j'}}\right)\right]_{(p+2)x(p+2)} = \sum_{i=1}^{m}\sum_{k=1}^{r}\left[x_{ji[1]k}x_{j'i[1]k}t_{i[1]k}\exp\left(-\mathbf{w}'_{i[1]k}\boldsymbol{\beta}\right)\right], \tag{17}$$

is a function of the double ranked auxiliary covariate used to draw the $\text{DERSS}_{\min}$. Fisher's information used to quantify the information or precision of the estimate of the vector of parameters. When making an inference about β_j, we need the Fisher's information of the risk factor coefficient, $I(\beta_j)$, to draw inference. Moreover, we assumed the ranking variable Z follows a distribution function denoted by $F_Z(z)$. We assume that Z has an absolutely continuous density $f_Z(z)$. Note that ranking on this variable stimulates some ordering on the response *T*, which leads to improved precision. Then the density function of $V_{(1)}$ is $g_{V_{(1)}} = mf_{Z_{(1)}}(z)[1-F_{Z_{(1)}}(z)]^{m-1} = m^2 f_Z(z)\,[1-F_Z(z)]^{m^2-1}$. Thus the information in an DERSS$_{\min}$ for the j^{th} parameter is

$$I(\beta_j)_{DERSS_{\min}} = -E_{V_{(1)}} E_{T_{[1]}|V_{(1)}} \left(\frac{\partial^2 L(\boldsymbol{\beta})}{\partial \beta_j^2} \right)$$

$$= E_{V(1)} E_{T_{[1]}|V(1)} \left(\sum_{i=1}^{m} \sum_{k=1}^{r} \left[x_{ji[1]k}^2 T_{i[1]k} \exp\left(-\mathbf{W}'_{i[1]k}\boldsymbol{\beta}\right) \right] \right)$$

$$= \sum_{i=1}^{m} \sum_{k=1}^{r} \int_z \left(\left[x_{ji[1]k}^2 E_{T_{[1]}|V_{(1)}}(T_{i[1]k}) \exp\left(-\mathbf{W}'_{i[1]k}\boldsymbol{\beta}\right) \right] \right) mf_{Z_{(1)}}(z)[1-F_{Z_{(1)}}(z)]^{m-1} dz$$

Since we are using DERSS$_{\min}$, and $\beta_{P+1} > 0$, which means { $\exp\left(-\mathbf{w}'_{i[1]k}\boldsymbol{\beta}\right)$ is a decreasing function of *z* and also $m[1-F_{Z(1)}(z)]^{m-1}$ is a decreasing function of *z* then by See and Chen (2008) we have

$$I(\beta_j)_{DERSS_{\min}}$$

$$\geq \sum_{i=1}^{m} \sum_{k=1}^{r} \int_Z \left(\left[x_{ji[1]k}^2 E_{T_{[1]}|V_{(1)}}(T_{i[1]k}) \exp\left(-\mathbf{W}'_{i[1]k}\boldsymbol{\beta}\right) \right] \right) f_{Z(1)}(z) dz \int_Z m f_{Z_{(1)}}(z)\,[1-F_{Z_{(1)}}(z)]^{m-1} dz$$

Since $\int_Z m f_{Z_{(1)}}(z)\,[1-F_{Z_{(1)}}(z)]^{m-1} dz = 1$, therefore,

$$I(\beta_j)_{DERSS_{\min}} \geq \sum_{i=1}^{m} \sum_{k=1}^{r} \int_Z \left(\left[x_{ji[1]k}^2 E_{T_{[1]}|V_{(1)}}(T_{i[1]k}) \exp\left(-\mathbf{w}'_{i[1]k}\boldsymbol{\beta}\right) \right] \right) f_{Z_{(1)}}(z) dz = I(\beta_j)_{ERSS_{\min}}. \quad (18)$$

However, Samawi et al. (2018) showed that $I(\beta_j)_{ERSS_{\min}} \geq I(\beta_j)_{SRS}$. Thus $I(\beta_j)_{DERSS_{\min}} \geq I(\beta_j)_{ERSS_{\min}} \geq I(\beta_j)_{SRS}$ Similarly, we can show that the inequality in (18) holds when using DERSS$_{\max}$ when the association between time to events and *Z* is negative ($\beta_{P+1} < 0$.)

AFT -Weibull Regression Model

As indicated by Samawi et al. (2018), "the Weibull distribution function ($W(\lambda,\delta)$) is usually formulated in the form of an extreme value distribution since log(T) can be expressed as a function of the Weibull parameters and follows extreme value distribution". Let T has $W(\lambda,\delta)$ distribution function, where λ is the scale parameter and δ is the shape parameter. Then the hazard function is given by

$$h_{ik}(t \mid \mathbf{w}_{i[1]k}, \boldsymbol{\beta}) = (\delta^*)^{-2} \exp\left(\frac{\log\ t_{i[1]k} - \mathbf{w}'_{i[1]k}\boldsymbol{\beta}}{\delta^*} \right),\ i = 1,2,...,m; k = 1,2,...,r. \quad (19)$$

Note that, the $\log\lambda$ is a coefficient, with the regression coefficient vector $\boldsymbol{\beta}$ and $\delta^* = \frac{1}{\delta}$.

Thus the survival function of *T* is

$$S_{ik}(t) = \exp\left[-\exp\left(\frac{\log\ t_{i[1]k} - \mathbf{w}'_{i[1]k}\boldsymbol{\beta}}{\delta^*} \right) \right], -\infty < \log t_{i[1]k} < \infty, \quad (20)$$

and the density function is

$$f(t_{ik}) = (\delta^*)^{-1}\exp\left[\frac{log\ t_{i[1]k} - \mathbf{w}'_{i[1]k}\boldsymbol{\beta}}{\delta^*} - \exp\left(\frac{log\ t_{i[1]k} - \mathbf{w}'_{[1]k}\boldsymbol{\beta}}{\delta^*}\right)\right], -\infty < \log t_{i[1]k} < \infty \,. \quad (21)$$

Then the log likelihood function can be simplified to

$$L(\boldsymbol{\beta}) = \log[l(\boldsymbol{\beta})] = \sum_{i=1}^{m}\sum_{k=1}^{r}\gamma_{ik}\left[(-\log\delta^*) + \left(\frac{log\ t_{i[1]k} - \mathbf{w}'_{i[1]k}\boldsymbol{\beta}}{\delta^*}\right)\right] - \exp\left(\frac{log\ t_{i[1]k} - \mathbf{w}'_{i[1]k}\boldsymbol{\beta}}{\delta^*}\right)$$

(22)

The MLE approach, for the j^{th} covariate can be obtained by solving

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \beta_j} = \sum_{i=1}^{r}\sum_{j=1}^{m}\left[-x_{ji[1]k}\{\gamma_{ik} - \exp\left(-\frac{log\ t_{i[1]k} - \mathbf{w}'_{i[1]k}\hat{\boldsymbol{\beta}}}{\delta^*}\right)\}\right] = 0. \quad (23)$$

Similar to the derivation ofthe exponential AFT model, we have, using the second partial derivative, the variance covariance matrix (Fisher information matrix) as

$$-\left[\left(\frac{\partial^2 L(\boldsymbol{\beta})}{\partial\beta_j\partial\beta_{j'}}\right)\right]_{(p+2)x(p+2)} = \sum_{i=1}^{m}\sum_{k=1}^{r}\left[\frac{x_{ji[1]k}x_{j'i[1]k}}{\delta^*}\exp\left(\frac{log\ t_{i[1]k} - \mathbf{w}'_{i[1]k}\boldsymbol{\beta}}{\delta^*}\right)\right]. \quad (24)$$

Again, (24) is a function of the double ranked auxiliary covariate. Similarly, as in exponential case, we can show that by See and Chen (2008) that

$$I(\beta_j)_{DERSS_{\min}} = -E_{V_{(1)}}E_{T_{[1]}|V_{(1)}}\left(\frac{\partial^2 L(\boldsymbol{\beta})}{\partial\beta_j^2}\right) \geq I(\beta_j)_{ERSS_{\min}}$$

$$\geq \sum_{i=1}^{m}\sum_{k=1}^{r}\int_Z\left(\frac{x^2_{ji[1]k}}{\delta^*}E_{T_{[1]}|V_{(1)}}\left[\exp\left((log\ t_{i[1]k} - \mathbf{w}'_{i[1]k}\boldsymbol{\beta})/\delta^*\right)\right]\right)f_Z(z)dz = I(\beta_j)_{SRS}$$

(25)

The inequality in (25) holds when the association between time to events and Z is negative ($\beta_{P+1} < 0$) and $DERSS_{max}$ is used.

AFT- Log-logistic Regression Model

In this section we provide the derivation for the log-logistic regression model. The hazard function is given by

$$h_{ik}(t\,|\,\mathbf{w}_{i[1]k},\boldsymbol{\beta}) = \frac{\exp\left(\frac{log\ t_{i[1]k} - \mathbf{w}'_{i[1]k}\boldsymbol{\beta}}{\lambda}\right)}{\lambda\left(1 + \frac{log\ t_{i[1]k} - \mathbf{w}'_{i[1]k}\boldsymbol{\beta}}{\lambda}\right)},\ i = 1,2,...,m; k = 1,2,...,r\,, \quad (26)$$

where λ is scale parameter for the log logistic distribution. Now using $DERSS_{min}$, the survival function for the log-logistic survival time T is given by:

$$S_{ik}(t) = \left[1 + \exp\left(\frac{log\ t_{i[1]k} - \mathbf{w}'_{i[1]k}\boldsymbol{\beta}}{\lambda}\right)\right]^{-1}, -\infty < \log t_{i[1]k} < \infty. \quad (27)$$

Similarly, the log likelihood function will simplify to

$$L(\boldsymbol{\beta})=\log[l(\boldsymbol{\beta})]=\sum_{i=1}^{m}\sum_{k=1}^{r}\gamma_{ik}\left\{(-\log\lambda)+\left(\frac{log\ t_{i[1]k}-\mathbf{w}'_{i[1]k}\boldsymbol{\beta}}{\lambda}\right)-2\log\left[1+\exp\left(\frac{log\ t_{i[1]k}-\mathbf{w}'_{i[1]k}\boldsymbol{\beta}}{\lambda}\right)\right]\right\}$$
$$-(1-\gamma_{ik})\log\left[1+\exp\left(\frac{log\ t_{i[1]k}-\mathbf{w}'_{i[1]k}\boldsymbol{\beta}}{\lambda}\right)\right] \quad (28)$$

The MLE estimates of the k^{th} parameter is obtained by solving

$$\frac{\partial L(\boldsymbol{\beta})}{\partial\beta_j}=\sum_{i=1}^{m}\sum_{k=1}^{r}\left[-x_{ji[1]k}\{\gamma_{ik}-\frac{\exp\left(\frac{log\ t_{i[1]k}-\mathbf{w}'_{i[1]k}\hat{\boldsymbol{\beta}}}{\lambda}\right)}{1+\exp\left(\frac{log\ t_{i[1]k}-\mathbf{w}'_{i[1]k}\hat{\boldsymbol{\beta}}}{\lambda}\right)}\}\right]+\frac{\gamma_{ik}x_{ji[1]k}\exp\left(\frac{log\ t_{i[1]k}-\mathbf{w}'_{i[1]k}\hat{\boldsymbol{\beta}}}{\lambda}\right)}{1+\exp\left(\frac{log\ t_{i[1]k}-\mathbf{w}'_{i[1]k}\hat{\boldsymbol{\beta}}}{\lambda}\right)}=0. \quad (29)$$

Also, the estimate of variance covariance matrix is given by:

$$-\left[\left(\frac{\partial^2 L(\boldsymbol{\beta})}{\partial\beta_j\partial\beta_{j'}}\right)\right]_{(p+2)x(p+2)}=\sum_{i=1}^{m}\sum_{k=1}^{r}\left[(1+\gamma_{ik})\frac{\frac{x_{ji[1]k}x_{j'i[1]k}}{\lambda^2}\exp\left(\frac{log\ t_{i[1]k}-\mathbf{w}'_{i[1]k}\hat{\boldsymbol{\beta}}}{\lambda}\right)}{\left[1+\exp\left(\frac{log\ t_{i[1]k}-\mathbf{w}'_{i[1]k}\hat{\boldsymbol{\beta}}}{\lambda}\right)\right]^2}\right]. \quad (30)$$

Similar to Samawi et al. (2018), let $D(v_{i(1)k})=\dfrac{\exp\left(\frac{log\ t_{i[1]k}-\mathbf{w}'_{i[1]k}\hat{\boldsymbol{\beta}}}{\lambda}\right)}{\left[1+\exp\left(\frac{log\ t_{i[1]k}-\mathbf{w}'_{i[1]k}\hat{\boldsymbol{\beta}}}{\lambda}\right)\right]^2}$. Note that $D(t_{i[1]k},v_{i(1)k})$ is a decreasing function in $v\le 0$, for $t\ge 0$. In practice, $v\le 0$ can be achieved by shifting *v* by a constant. Thus,

$$I(\beta_j)_{DERSS_{\min}}=E_{V_{(1)}}E_{T_{[1]}|V_{(1)}}\left(\sum_{i=1}^{m}\sum_{k=1}^{r}\left[(1+\gamma_{ik})\frac{x^2_{ji[1]k}}{\lambda^2}D(T_{i[1]k},V_{i(1)k})\right]\right)$$
$$=E_{V_{(1)}}\left(\sum_{i=1}^{m}\sum_{k=1}^{r}\left[(1+\gamma_{ik})\frac{x^2_{ji[1]k}}{\lambda^2}E_{T_{[1]}|V_{(1)}}(D(T_{i[1]k},V_{i(1)k}))\right]\right)$$
$$=\sum_{i=1}^{m}\sum_{j=1}^{r}\left[(1+\gamma_{ik})\frac{x^2_{ji[1]k}}{\lambda^2}E_{V_{(1)}}\left\{E_{T_{[1]}|V_{(1)}}(D(T_{i[1]k},V_{i(1)k}))\right\}\right] \quad (31)$$
$$=\sum_{i=1}^{m}\sum_{k=1}^{r}\left[(1+\gamma_{ik})\frac{x^2_{ji[1]k}}{\lambda^2}\int_v E_{T_{[1]}|V_{(1)}}(D(T_{i[1]k},v))m[1-F_{Z_{(1)}}(v)]^{m-1}f_{Z_{(1)}}(v)dv\right]$$

Moreover, $m(1-F_{Z(1)}(v))^{m-1}$is a decreasing function of *v*; therefore, by See and Chen (2008)

$$I(\beta_j)_{DERSS_{\min}}\ge\sum_{i=1}^{m}\sum_{k=1}^{r}\left[(1+\gamma_{ik})\frac{x^2_{ji[1]k}}{\lambda^2}\int_v E_{T_{[1]}|V_{(1)}}(D(T_{i[1]k},v))f_{Z_{(1)}}(v)dv\int_v m[1-F_{Z_{(1)}}(v)]^{m-1}f_{Z_{(1)}}(v)dv\right] \quad (32)$$

where, $\int_v m[1-F_{Z_{(1)}}(v)]^{m-1} f_{Z_{(1)}}(v)dv = 1$. This implies

$$I(\beta_j)_{DERSS_{\min}} \geq \sum_{i=1}^{m}\sum_{k=1}^{r}\left[\int_v (1+\gamma_{ik})\frac{x_{ji[1]k}^2}{\lambda^2} E_{T_{[1]}|V_{(1)}}(D(T_{i[1]k},v)) f_{Z(1)}(v)dv\right] = I(\beta_j)_{ERSS_{\min}} \geq I(\beta_j)_{SRS} \quad (33)$$

where the last inequality is by Samawi et al. (2018). Similarly, we can show that the inequality in (33) holds for $\beta_{P+1} < 0$ and the association between time to events and *Z* is negative when $\text{DERSS}_{\max}$ is used. Finally, the other AFT regression model, including log-normal and Gamma, will have similar derivations and their performance will be discussed next in the simulation section. Finally, comparing with SRS, ERSS and DERSS provide larger percentage of events and require smaller sample sizes.

Simulation Studies

Simulation studies are designed to get insight the performance of the AFT models using $\text{DERSS}_{\min}$ compared with $\text{ERSS}_{\min}$ and SRS. The performance of all AFT models with respect to the power of hypotheses testing and parameter estimations is discussed. As in Samawi et al. (2018), we consider values of the conditional hazards ratios range from 1 to 1.649, and the associations between survival time and the auxiliary covariate are 0.2 and 0.5. The AFT models considered are the exponential, Weibull, log-logistic, log-normal and Gamma. The set sizes used are *m*=10, 15 and cycle size is *r*=15. We repeated the process 5000 times to compute the accuracy of AFT models performance.

In our simulations, β_1 represents the parameter associated with the auxiliary covariate (*Z*), β_2 is the parameter associated with the risk factor of interest (*X*) and $\sigma = 1$. However, the other parameter involved in the simulation are given in the tables due to the nature of the underlying distribution. The empirical nominal value, $\alpha = 0.05$, is considered.

Table 1 and 4 show that for testing the hypothesis $H_o : \beta_2 = 0$ vs $H_a : \beta_2 \neq 0$, when controlling for the ranked auxiliary covariate, the $\text{DERSS}_{\min}$ results in a more powerful test comparingwith using $\text{ERSS}_{\min}$ and SRS.

All sampling schemed considered in this simulation, SRS , $\text{ERSS}_{\min}$ and $\text{DERSS}_{\min}$, achieved close estimation to the test nominal value (0.05) under the null hypothesis. The simulation results indicate that the power of the test is a monotone increasing function of the set size *m* increase and/or the value of β_1. In addition, the simulation results indicate that using $\text{DERSS}_{\min}$ for the AFT models provide greater power than both $\text{ERSS}_{\min}$ and SRS, in all cases. Note that the relative efficiency based on the power of the test between $\text{DERSS}_{\min}$ and $\text{ERSS}_{\min}$ ranges from 1.02-1.80 and between $\text{DERSS}_{\min}$ and SRS ranges from 1.5-2.5. Finally, Table 2, 3, 5, and 6, demonstrate that $\text{DERSS}_{\min}$ provides more efficient estimators of the hazards ratios in terms of smaller MSEs and bias as well as narrower confidence intervals for all parametric models comparing with ERSS and SRS.

Table 1: Estimation of $(\alpha = 0.05)$ **and the power of testing** $H_o : \beta_2 = 0$ vs $H_a : \beta_2 \neq 0$ **adjusting for the auxiliary variable (*Z*) in the model. {Censoring variable= U(0,1)*2}}**

		(Weibull $\delta^* = 1$**)(Exponential) , *m=10 and r=15***					
		DERSSmin		**ERSSmin**		**SRS**	
β_1	β_2	**Events %**	**The Power Function**	**Events %**	**The Power Function**	**Events %**	**The Power Function**
0.2	0.0	42	0.051	36	0.052	29	0.054
0.2	0.2	45	0.349	39	0.326	30	0.261
0.2	0.5	50	0.984	42	0.968	31	0.900
0.5	0.0	64	0.052	50	0.049	31	0.054
0.5	0.2	66	0.487	52	0.408	31	0.262
0.5	0.5	69	0.996	54	0.989	32	0.913
(Weibull $\delta^* = 1$**) (Exponential) *m=15 and r=15***							
0.2	0.0	44	0.050	38	0.049	30	0.049
0.2	0.2	47	0.530	40	0.448	30	0.367
0.2	0.5	52	0.999	44	0.998	31	0.980
0.5	0.0	68	0.054	53	0.052	31	0.055
0.5	0.2	70	0.683	55	0.583	31	0.375
0.5	0.5	73	1.000	57	0.999	32	0.980
(Weibull $\delta^* = 1.5$**), *m=10 and r=15***							
0.2	0.0	33	0.054	26	0.051	16	0.044
0.2	0.2	38	0.835	29	0.709	17	0.475
0.2	0.5	45	1.000	34	1.000	20	0.998
0.5	0.0	65	0.060	45	0.053	19	0.050
0.5	0.2	68	0.967	48	0.905	20	0.543
0.5	0.5	70	1.000	51	1.000	22	0.998
(Weibull $\delta^* = 1.5$**), *m=15 and r=15***							
0.2	0.0	36	0.056	27	0.052	16	0.055
0.2	0.2	41	0.960	31	0.892	17	0.659
0.2	0.5	48	1.000	36	1.000	20	1.000
0.5	0.0	70	0.056	50	0.054	19	0.052
0.5	0.2	72	0.997	52	0.985	20	0.722
0.5	0.5	74	1.000	54	1.000	23	1.000
(**Gamma** $\lambda = 1.5$**), *m=15 and r=20***							
0.2	0.0	51	0.057	45	0.058	37	0.055
0.2	0.2	58	0.520	50	0.443	38	0.314
0.2	0.5	69	0.995	57	0.990	39	0.952
0.5	0.0	73	0.06	59	0.058	38	0.059
0.5	0.2	79	0.698	64	0.558	39	0.323
0.5	0.5	85	1.000	69	0.995	39	0.956

Table 2: Hazard ratio (HR) estimation {Censoring variable=U(0,1)*2}

		m=10 and r=15 **(Weibull $\delta^*=1$) (Exponential)**								
		DERSS$_{min}$			**ERSS$_{min}$**			**SRS**		
β_1	**HR**	**Estimate**	**MSE**	**Bias**	**Estimate**	**MSE**	**Bias**	**Estimate**	**MSE**	**Bias**
0.2	1.000	1.006	0.019	0.006	1.010	0.022	0.010	1.014	0.028	0.014
0.2	1.221	1.236	0.027	0.015	1.241	0.033	0.020	1.244	0.042	0.023
0.2	1.649	1.681	0.048	0.032	1.685	0.061	0.036	1.695	0.085	0.046
0.5	1.000	1.008	0.012	0.009	1.006	0.015	0.006	1.015	0.026	0.015
0.5	1.221	1.230	0.018	0.008	1.237	0.024	0.015	1.243	0.040	0.021
0.5	1.649	1.671	0.033	0.022	1.676	0.044	0.027	1.690	0.079	0.041
m=15 and r=15 **(Weibull $\delta^*=1$)(Exponential)**										
0.2	1.000	1.007	0.012	0.007	1.005	0.013	0.005	1.009	0.018	0.009
0.2	1.221	1.233	0.017	0.012	1.230	0.020	0.090	1.235	0.026	0.014
0.2	1.649	1.667	0.030	0.018	1.672	0.034	0.023	1.677	0.047	0.028
0.5	1.000	1.003	0.007	0.003	1.002	0.009	0.002	1.008	0.017	0.008
0.5	1.221	1.226	0.011	0.005	1.231	0.014	0.010	1.235	0.026	0.013
0.5	1.649	1.656	0.019	0.007	1.663	0.025	0.014	1.675	0.047	0.026
m=10 and r=15 **(Weibull $\delta^*=1.5$)**										
0.2	1.000	1.002	0.006	0.002	1.002	0.008	0.002	1.008	0.013	0.008
0.2	1.221	1.230	0.008	0.008	1.230	0.012	0.011	1.235	0.021	0.014
0.2	1.649	1.662	0.016	0.014	1.665	0.022	0.016	1.680	0.045	0.031
0.5	1.000	1.004	0.003	0.004	1.004	0.004	0.004	1.008	0.011	0.008
0.5	1.221	1.225	0.004	0.004	1.229	0.006	0.008	1.231	0.017	0.010
0.5	1.649	1.656	0.008	0.007	1.658	0.013	0.009	1.672	0.037	0.024
m=15 and r=15 **(Weibull $\delta^*=1.5$)**										
0.2	1.000	1.001	0.004	0.001	1.004	0.005	0.004	1.002	0.008	0.002
0.2	1.221	1.227	0.005	0.006	1.227	0.007	0.005	1.231	0.013	0.010
0.2	1.649	1.655	0.009	0.006	1.657	0.013	0.008	1.668	0.029	0.020
0.5	1.000	1.001	0.002	0.001	1.000	0.003	0.000	1.001	0.007	0.002
0.5	1.221	1.223	0.003	0.002	1.224	0.004	0.003	1.230	0.011	0.008
0.5	1.649	1.652	0.005	0.003	1.654	0.008	0.005	1.663	0.023	0.014
(Gamma $\lambda=1.5$), ***m=10 and r=15***										
0.2	1.000	1.008	0.013	0.008	1.010	0.016	0.010	1.010	0.022	0.010
0.2	1.221	1.229	0.017	0.008	1.231	0.021	0.010	1.242	0.033	0.021
0.2	1.649	1.660	0.025	0.011	1.655	0.036	0.007	1.674	0.064	0.026
0.5	1.000	1.004	0.008	0.004	1.005	0.010	0.005	1.010	0.021	0.010
0.5	1.221	1.228	0.011	0.007	1.227	0.015	0.006	1.233	0.029	0.011
0.5	1.649	1.657	0.018	0.008	1.662	0.026	0.014	1.663	0.058	0.015

Table 3: 95% Confidence Interval of the Hazard Ratio (HR) {Censoring variable=U(0,1)*2}

		m=10 and r=15 **(Weibull $\delta^*=1$) (Exponential)**								
		DERSS$_{min}$			**ERSS$_{min}$**			**SRS**		
β_1	**HR**	**Lower**	**Upper**	**Coverage Probability**	**Lower**	**Upper**	**Coverage Probability**	**Lower**	**Upper**	**Coverage Probability**
0.2	1.000	0.778	1.302	0.949	0.767	1.331	0.948	0.745	1.382	0.946
0.2	1.221	0.962	1.589	0.948	0.947	1.628	0.944	0.914	1.696	0.950
0.2	1.649	1.314	2.150	0.946	1.288	2.204	0.951	1.240	2.318	0.945
0.5	1.000	0.819	1.243	0.947	0.794	1.275	0.950	0.749	1.376	0.946
0.5	1.221	1.001	1.512	0.943	0.979	1.563	0.943	0.918	1.685	0.949
0.5	1.649	1.358	2.056	0.946	1.324	2.121	0.946	1.244	2.298	0.944
m=15 and r=15 **(Weibull $\delta^*=1$)(Exponential)**										
0.2	1.000	0.821	1.236	0.950	0.807	1.252	0.951	0.786	1.295	0.951
0.2	1.221	1.011	1.510	0.943	0.992	1.527	0.937	0.962	1.586	0.954
0.2	1.649	1.372	2.026	0.945	1.351	2.070	0.951	1.304	2.157	0.953
0.5	1.000	0.851	1.183	0.946	0.832	1.210	0.948	0.788	1.289	0.945
0.5	1.221	1.042	1.442	0.946	1.024	1.480	0.949	0.966	1.579	0.950
0.5	1.649	1.407	1.950	0.953	1.381	2.001	0.950	1.308	2.147	0.947
m=10 and r=15 **(Weibull $\delta^*=1.5$)**										
0.2	1.000	0.867	1.158	0.946	0.849	1.184	0.949	0.816	1.249	0.956
0.2	1.221	1.069	1.416	0.946	1.046	1.453	0.944	0.994	1.538	0.947
0.2	1.649	1.441	1.918	0.940	1.406	1.974	0.942	1.334	2.121	0.946
0.5	1.000	0.906	1.113	0.940	0.887	1.136	0.946	0.831	1.225	0.950
0.5	1.221	1.106	1.356	0.945	1.086	1.392	0.946	1.012	1.502	0.950
0.5	1.649	1.490	1.840	0.938	1.456	1.888	0.940	1.355	2.068	0.941
m=15 and r=15 **(Weibull $\delta^*=1.5$)**										
0.2	1.000	0.894	1.121	0.944	0.881	1.144	0.948	0.846	1.188	0.945
0.2	1.221	1.100	1.369	0.943	1.08	1.395	0.945	1.034	1.467	0.947
0.2	1.649	1.451	1.848	0.948	1.452	1.891	0.946	1.386	2.011	0.943
0.5	1.000	0.923	1.085	0.944	0.909	1.101	0.946	0.857	1.171	0.948
0.5	1.221	1.129	1.330	0.946	1.112	1.348	0.947	1.049	1.443	0.947
0.5	1.649	1.521	1.793	0.944	1.496	1.829	0.944	1.405	1.971	0.947
(Gamma $\lambda=1.5$), ***m=10 and r=15***										
0.2	1.000	0.812	1.252	0.943	0.798	1.280	0.9424	0.769	1.329	0.945
0.2	1.221	1.008	1.501	0.936	0.986	1.539	0.948	0.946	1.636	0.943
0.2	1.649	1.375	1.998	0.943	1.332	2.100	0.930	1.261	2.240	0.943
0.5	1.000	0.849	1.195	0.938	0.830	1.220	0.941	0.776	1.317	0.941
0.5	1.221	1.046	1.50	0.944	1.018	1.588	0.943	0.948	1.608	0.947
0.5	1.649	1.421	1.933	0.938	1.382	2.010	0.938	1.271	2.184	0.929

Table 4: Estimation of $(\alpha = 0.05)$ **and the power of testing** $H_o : \beta_2 = 0$ vs $H_a : \beta_2 \neq 0$ **adjusting for the auxiliary variable (*Z*) in the model. {Censoring variable= U(0,1)*2}.**

		(Log-logistic $\lambda = 1.5$**), *m=10 and r=15***					
		DERSSmin		**ERSSmin**		**SRS**	
β_1	β_2	**Events %**	**The Power Function**	**Events %**	**The Power Function**	**Events %**	**The Power Function**
0.2	0.0	31	0.050	26	0.051	19	0.049
0.2	0.2	31	0.352	26	0.331	19	0.256
0.2	0.5	32	0.975	27	0.956	20	0.908
0.5	0.0	52	0.055	38	0.048	20	0.043
0.5	0.2	52	0.466	38	0.396	20	0.257
0.5	0.5	52	0.994	39	0.980	21	0.906
(Log-logistic $\lambda = 1.5$**), *m=15 and r=15***							
0.2	0.0	32	0.047	26	0.047	19	0.045
0.2	0.2	32	0.521	26	0.460	19	0.379
0.2	0.5	33	0.999	28	0.996	20	0.985
0.5	0.0	56	0.050	41	0.047	20	0.050
0.5	0.2	56	0.639	41	0.58	20	0.381
0.5	0.5	55	1.000	41	0.999	22	0.984
(Log-Normal $\sigma = 1.5$**), *m=10 and r=15***							
0.2	0.0	34	0.050	30	0.048	24	0.044
0.2	0.2	36	0.454	31	0.427	24	0.379
0.2	0.5	40	0.997	33	0.992	25	0.982
0.5	0.0	51	0.048	40	0.054	24	0.047
0.5	0.2	53	0.520	41	0.465	25	0.372
0.5	0.5	56	0.999	43	0.997	25	0.981
(Log-Normal $\sigma = 1.5$**), *m=15 and r=15***							
0.2	0.0	35	0.051	30	0.050	24	0.050
0.2	0.2	37	0.620	32	0.580	24	0.512
0.2	0.5	41	1.000	35	1.000	25	0.999
0.5	0.0	54	0.049	42	0.053	24	0.049
0.5	0.2	56	0.690	43	0.642	25	0.511
0.5	0.5	59	1.000	46	1.000	26	0.998

Table 5: Hazard ratio (HR) estimation {Censoring variable=U(0,1)*2}

		m=10 and r=15, **(Log-logistic $\lambda = 1.5$)**								
		DERSS$_{min}$			**ERSS$_{min}$**			**SRS**		
β_1	**HR**	**Estimate**	**MSE**	**Bias**	**Estimate**	**MSE**	**Bias**	**Estimate**	**MSE**	**Bias**
0.2	1.000	1.006	0.017	0.006	1.005	0.019	0.005	1.013	0.026	0.013
0.2	1.221	1.232	0.026	0.011	1.240	0.031	0.019	1.242	0.040	0.021
0.2	1.649	1.675	0.058	0.026	1.674	0.065	0.025	1.688	0.093	0.040
0.5	1.000	1.009	0.013	0.010	1.006	0.015	0.006	1.014	0.024	0.014
0.5	1.221	1.229	0.018	0.008	1.231	0.023	0.009	1.236	0.038	0.014
0.5	1.649	1.660	0.037	0.012	1.667	0.048	0.018	1.679	0.085	0.030
m=15 and r=15, **(Log-logistic $\lambda = 1.5$)**										
0.2	1.000	1.005	0.010	0.005	1.005	0.012	0.005	1.009	0.015	0.009
0.2	1.221	1.231	0.016	0.009	1.229	0.018	0.008	1.237	0.025	0.016
0.2	1.649	1.666	0.036	0.018	1.663	0.041	0.014	1.672	0.053	0.023
0.5	1.000	1.004	0.008	0.004	1.002	0.010	0.002	1.010	0.016	0.010
0.5	1.221	1.225	0.011	0.003	1.226	0.014	0.004	1.235	0.024	0.013
0.5	1.649	1.659	0.023	0.010	1.660	0.029	0.011	1.674	0.053	0.025
m=10 and r=15 **(Log-Normal $\sigma = 1.5$)**										
0.2	1.000	1.003	0.012	0.003	1.007	0.013	0.007	1.007	0.015	0.007
0.2	1.221	1.230	0.019	0.009	1.232	0.021	0.011	1.234	0.024	0.012
0.2	1.649	1.656	0.035	0.007	1.662	0.038	0.014	1.666	0.048	0.018
0.5	1.000	1.007	0.011	0.007	1.005	0.012	0.005	1.007	0.016	0.007
0.5	1.221	1.230	0.016	0.009	1.230	0.018	0.006	1.231	0.025	0.009
0.5	1.649	1.656	0.028	0.008	1.663	0.033	0.014	1.663	0.048	0.014
m=15 and r=15 **(Log-Normal $\sigma = 1.5$)**										
0.2	1.000	1.003	0.008	0.003	1.005	0.009	0.005	1.004	0.010	0.004
0.2	1.221	1.227	0.012	0.006	1.226	0.014	0.005	1.227	0.015	0.006
0.2	1.649	1.656	0.023	0.007	1.660	0.025	0.011	1.657	0.031	0.009
0.5	1.000	1.003	0.007	0.003	1.004	0.008	0.004	1.007	0.010	0.007
0.5	1.221	1.223	0.010	0.002	1.225	0.011	0.004	1.226	0.016	0.004
0.5	1.649	1.655	0.018	0.006	1.656	0.022	0.008	1.656	0.030	0.009

Table 6: 95% confidence Interval of the Hazard Ratio (HR) {Censoring variable=U(0,1)*2}

		m=10 and r=15 **(Log-logistic $\lambda = 1.5$)**								
		$DERSS_{min}$			**$ERSS_{min}$**			**SRS**		
β_1	**HR**	**Lower**	**Upper**	**Coverage Probability**	**Lower**	**Upper**	**Coverage Probability**	**Lower**	**Upper**	**Coverage Probability**
0.2	1.000	0.789	1.285	0.950	0.776	1.305	0.949	0.755	1.362	0.951
0.2	1.221	0.960	1.583	0.946	0.950	1.622	0.951	0.919	1.686	0.954
0.2	1.649	1.282	2.192	0.948	1.257	2.233	0.944	1.219	2.348	0.943
0.5	1.000	0.819	1.245	0.945	0.801	1.266	0.952	0.759	1.359	0.957
0.5	1.221	0.995	1.519	0.946	0.976	1.554	0.945	0.917	1.670	0.950
0.5	1.649	1.331	2.073	0.944	1.302	2.137	0.939	1.219	2.321	0.944
m=15 and r=15 **(Log-logistic $\lambda = 1.5$)**										
0.2	1.000	0.827	1.222	0.953	0.817	1.239	0.953	0.797	1.281	0.955
0.2	1.221	1.009	1.503	0.954	0.994	1.522	0.957	0.969	1.581	0.952
0.2	1.649	1.346	2.065	0.941	1.323	2.092	0.946	1.286	2.179	0.949
0.5	1.000	0.849	1.188	0.949	0.835	1.202	0.953	0.798	1.280	0.950
0.5	1.221	1.034	1.451	0.951	1.019	1.474	0.949	0.975	1.571	0.955
0.5	1.649	1.391	1.979	0.948	1.365	2.020	0.946	1.292	2.172	0.951
m=10 and r=15 **(Log-Normal $\sigma = 1.5$)**										
0.2	1.000	0.807	1.248	0.950	0.804	1.262	0.952	0.794	1.280	0.957
0.2	1.221	0.992	1.526	0.950	0.985	1.543	0.947	0.969	1.572	0.952
0.2	1.649	1.333	2.058	0.949	1.323	2.090	0.950	1.295	2.146	0.947
0.5	1.000	0.827	1.226	0.952	0.816	1.238	0.946	0.793	1.279	0.953
0.5	1.221	1.011	1.497	0.944	0.996	1.513	0.948	0.966	1.569	0.947
0.5	1.649	1.359	2.019	0.948	1.344	2.057	0.948	1.292	2.142	0.951
m=15 and r=15 **(Log-Normal $\sigma = 1.5$)**										
0.2	1.000	0.842	1.195	0.949	0.838	1.206	0.950	0.826	1.220	0.950
0.2	1.221	1.031	1.460	0.948	1.023	1.470	0.947	1.008	1.495	0.955
0.2	1.649	1.390	1.972	0.947	1.381	1.996	0.947	1.350	2.035	0.947
0.5	1.000	0.856	1.176	0.950	0.849	1.188	0.947	0.829	1.223	0.951
0.5	1.221	1.045	1.432	0.942	1.036	1.449	0.948	1.006	1.498	0.948
0.5	1.649	1.412	1.939	0.953	1.397	1.964	0.945	1.349	2.033	0.955

Illustration based Worcester Heart Attack Study

Age at baseline is an important covariate in all survival studies. Due to the availability of baseline age variable, age can be considered as an auxiliary variable for subject's selection in survival studies. For illustration purposes, Worcester Heart Attack Study data (Hosmer et al., 2008) is used. In this section, we illustrate the use of DERSS, ERSS and SRS for the AFT survival models.

Worcester Heart Attack Study consists of 500 subjects. The study investigates some factors, such as age, initial heart rate and BMI, that may influence survival time after a heart attack. The follow up time for all subjects in the study is initiated at the time of hospital admission after a heart attack and ends with death or loss to follow up (censoring). The variables used in the illustration are:

- Lenfol: The length of follow up, terminated either by death or censoring.

- Fstat: The censoring variable, loss to follow up=0, death=1.
- Age: Age at hospitalization.
- BMI: body mass index.
- HR: Initial heart rate

In the study the data are subjected to right-censoring only.

For illustrations purposes, we used the whole data as a population. Then we randomly draw $DERSS_{max}$, $ERSS_{max}$ and SRS samples of size n=100 (m=10, r=10) each. We used the auxiliary variable age as the ranking variable to $DERSS_{max}$ and $ERSS_{max}$ because of the negative association between age and survival time. The hypotheses of interest for this study are whether or not the baseline age, BMI and HR risk factors have effects on the length of survival time after a heart attack. We found that AFT Weibull model best fits the data. Table 7 is the results of the AFT survival analysis based on the whole data (N=500) using the Weibull model.

In addition, Table 7 provides the AFT model analysis with and without age. From table 7, the association between BMI and time to death is positive indicating that BMI is a protective factor controlling for Age and HR. However, the negative association of HR indicates that HR is a risk factor for survival time controlling for age and BMI.

Table 7: AFT-Weibull model analysis using all the data

Analysis of Maximum Likelihood Parameter Estimates							
Parameter	**DF**	**Estimate**	**Standard Error**	**95% Confidence Limits**		**Chi-Square**	**Pr > ChiSq**
Intercept	1	16.2112	1.5578	13.1580	19.2644	108.30	<.0001
AGE	1	-0.1082	0.0135	-0.1347	-0.0818	64.46	<.0001
BMI	1	0.0798	0.0300	0.0210	0.1386	7.08	0.0078
HR	1	-0.0248	0.0054	-0.0353	-0.0143	21.49	<.0001
Scale	1	1.9351	0.1155	1.7214	2.1754		
Weibull Shape	1	0.5168	0.0309	0.4597	0.5809		
Using Weibull model without age in the model							
Parameter	**DF**	**Estimate**	**Standard Error**	**95% Confidence Limits**		**Chi-Square**	**Pr > ChiSq**
Intercept	1	6.0349	0.9202	4.2313	7.8384	43.01	<.0001
BMI	1	0.1915	0.0311	0.1305	0.2525	37.87	<.0001
HR	1	-0.0311	0.0056	-0.0420	-0.0202	31.16	<.0001
Scale	1	2.0388	0.1242	1.8093	2.2974		
Weibull Shape	1	0.4905	0.0299	0.4353	0.5527		

Table 8: AFT model (Weibull) using $DERSS_{max}$ without age in the model

Analysis of Maximum Likelihood Parameter Estimates							
Parameter	**DF**	**Estimate**	**Standard Error**	**95% Confidence Limits**		**Chi-Square**	**Pr > ChiSq**
Intercept	1	5.9498	1.0845	3.8242	8.0755	30.10	<.0001
BMI	1	0.2267	0.0559	0.1172	0.3362	16.47	<.0001
HR	1	-0.0209	0.0106	-0.0417	-0.0000	3.85	0.0496
Scale	1	1.8494	0.1571	1.5658	2.1844		
Weibull Shape	1	0.5407	0.0459	0.4578	0.6387		

Table 9: AFT model (Weibull) using $ERSS_{max}$ without age in the model

Analysis of Maximum Likelihood Parameter Estimates							
Parameter	**DF**	**Estimate**	**Standard Error**	**95% Confidence Limits**		**Chi-Square**	**Pr > ChiSq**
Intercept	1	2.5322	1.5315	-0.4695	5.5339	2.73	0.0983
BMI	1	0.0665	0.0478	-0.0273	0.1603	1.93	0.1645
HR	1	-0.0175	0.0104	-0.0379	0.0029	2.83	0.0926
Scale	1	1.8900	0.1804	1.5676	2.2787		
Weibull Shape	1	0.5291	0.0505	0.4388	0.6379		

Table 8 provides the result of AFT survival analysis of the DERSSmax sample of size m=5, r=20 (n=100). Table 9 shows the AFT analysis when using ERSSmax sample of size m=5, r=20 (n=100). Table 10 shows the analysis of AFT model using SRS of size 100. For this illustration DERSSmax provides the closest analysis to the whole data.

Table 10: AFT model (Weibull) using SRS without age in the model

Analysis of Maximum Likelihood Parameter Estimates							
Parameter	**DF**	**Estimate**	**Standard Error**	**95% Confidence Limits**		**Chi-Square**	**Pr > ChiSq**
Intercept	1	4.8184	1.9941	0.9100	8.7268	5.84	0.0157
BMI	1	0.1452	0.0603	0.0270	0.2634	5.80	0.0160
HR	1	-0.0066	0.0125	-0.0311	0.0180	0.27	0.6001
Scale	1	1.6124	0.2210	1.2325	2.1093		
Weibull Shape	1	0.6202	0.0850	0.4741	0.8113		

Final remarks

DERSS is a cost effective and efficient sampling technique compared with SRS and ERSS. In this paper we proposed a more efficient survival regression analysis method for AFT models based on the modified DERSS with ranking based on an auxiliary variable known to be associated with the response variable. We studied parameters estimation based on the maximum likelihood approach and provided an expression for the estimated variance-covariance matrix based on the inverse information matrix. The asymptotic behavior of the ML estimators was discussed. We concluded that using the modified DERSS can result in significant increase in power when implemented in the AFT models. Our simulation studies showed that in general, the power of the test increases as the set size *m* increases. In addition, DERSS provides more efficient inference of the parameters associated with hazard rates which results in smaller MSEs and narrower confidence intervals than those under SRS and ERSS.

References

1. Agresti, A. (2002). *An introduction to categorical data analysis* (Vol. 135): Wiley New York.
2. Al-Saleh, M. F., and Samawi, H. M. (2000). On the efficiency of Monte Carlo methods using steady state ranked simulated samples. Communications in Statistics-Simulation and Computation, 29(3), 941-954.
3. Al-Saleh, M. F., and Zheng, G. (2003). Controlled sampling using ranked set sampling. Journal of Nonparametric Statistics, 15(4-5), 505-516.
4. Albert, A., & Anderson, J. (1984). On the existence of maximum likelihood estimates in logistic regression models. Biometrika, 71(1), 1-10.
5. Chapman, J. W., Trudeau, M. E., Pritchard, K. I., Sawka, C. A., Mobbs, B. G., Hanna, W. M., Kahn, H., McCready, D. R., Lickley, L. A., (1992). A comparison of all-subset Cox and accelerated failure time models with Cox step-wise regression for node-positive breast cancer. Breast Cancer Research and Treatment, 22(3), 263 – 272.
6. Chen, Z. (2007). Ranked set sampling: Its essence and new applications. Environ. Ecol. Stat. 14,355–363.
7. Chen, Z., Zheng, G., Ghosh, K., Li, Z.(2005). Linkage disequilibrium mapping of quantitative trait loci by selective genotyping. Am. J. Hum. Genet. 77, 661–669.
8. Collett, D., 2003. Modelling Survival Data in Medical Research, second ed. Chapman and Hall, New York
9. Cox, D. R. (1972). Regression models and life-table. Journal of the Royal Statistical Society. Series B (Methodological), 34 (2), 187-220.

10. Hosmer, DW, Lemeshow, S, May S. (2008). *Applied Survival Analysis.* John Wiley & Sons, Inc. Hoboken.
11. Jabrah, R., Samawi, H., Vogel, R., Rochani, H., Linder, D. and Klibert, J. (2017). Using ranked auxiliary covariate as a more efficient sampling design for ANCOVA model: Analysis of a psychological intervention to buttress resilience. Communications for Statistical Applications and Methods, 24(3), 241-257.
12. Lehmann, E. L. and Casella,G. (1998). Theory of Point Estimation. Second, Ed. Springer, New York.
13. Liu, X. (2012). Survival analysis, models and applications. A John Wiley & Sons, Ltd., New York
14. McIntyre, G.A. (1952). A method for unbiased selective sampling, using ranked sets. Australian Journal of Agriculture Research, 3, 385-390.
15. Rashid, M., and Shifa, N., (2009). Consistency of the maximum likelihood estimator in logistic regression model: A different approach. Journal of Statistics 16, 1-11.
16. Rochani, H., Linder, D. F., Samawi, H., & Panchal, V. (2018). On inference of multivariate means under ranked set sampling. Communications for Statistical Applications and Methods, 25(1), 1-13.
17. Royston, P. (2001).The lognormal distribution as a model for survival time in cancer, with anemphasis on prognostic factors. Statistica Neerlandica. 55,89-104.
18. Samawi, H. M. (2002). On Double Extreme Ranked Set Sample with Application to Regression Estimator. Metron, LX(1-2), 53-66.
19. Samawi, H., and Ababneh, F. (2001). On regression analysis using ranked set sample. Journal of statistical Research (JSR), 35(2), 93-105.
20. Samawi, H. M., and Abu-Dayyeh, W., (2002). On regression analysis with random regressors using ranked samples. International journal of information and management sciences, 13(3), 19-36.
21. Samawi, H. M., Ahmed, M. S., and Abu-Dayyeh, W. (1996). Estimating the population mean using extreme ranked set sampling. Biometrical Journal, 38(5), 577-586.
22. Samawi, H. M. and Al-Sageer, O. A. (2001). On the estimation of the distribution function using extreme and median ranked set sampling. *Biom. J.* **43**(3), 357-373.
23. Samawi, H. M., and Al-Saleh, M. F. (2002). On regression analysis using bivariate ranked set samples. Metron, 60(3), 29-48.
24. Samawi, H. M., Al-Saleh, M. F., and Al-Saidy, O. (2009). The matched pair sign test using bivariate ranked set sampling for different ranking based schemes. Statistical Methodology, 6(4), 397-407.
25. Samawi, H. M., Helu, A., Rochani, H., Yin, J., Yu, L., & Vogel, R. (2018). Reducing sample size needed for accelerated failure time model using more efficient sampling methods. Journal of Statistical Theory and Practice, 12(3), 530-541
26. Scheaffer, R. L., Ill, W. M., and Ott, R. L. (2006). Elementary Survey Sampling. 6th edition, Thomson, Brooks/Cole, USA.
27. See C and Chen J. (2008). Inequalities on the variance of convex functions of random variables. Journal of inequalities in pure and applied mathematics. 9(3):1–5.
28. Sinsheimer, J.S., Blangero, J., Lange, K.(2000). Gamete competition models. Am. J. Hum. Genet. 66, 1168–1172.
29. Swindell, W. R. (2009). Accelerated failure time models provide a useful statsitical framework for aging research. Exp. Gerontol. 44(3): 190–200. doi:10.1016/j.exger.2008.10.005.
30. Zheng, G., Ghosh, K., Chen, Z., Li, Z.(2006). Extreme rank selections for linkage analysis of quantitative trait loci using selected sibpairs. Ann. Hum. Genet. 70, 857–866.