

On Optimal Designs of Some Censoring Schemes

Adnan Mohammad Awad
Department of Mathematics, University of Jordan
Jordan
adnanawad1947@gmail.com

Abstract

The main objective of this paper is to explore suitability of some entropy-information measures for introducing a new optimality censoring criterion and to apply it to some censoring schemes from some underlying life-time models. In addition, the paper investigates four related issues namely; the effect of the parameter of parent distribution on optimal scheme, equivalence of schemes based on Shannon and Awad sup-entropy measures, the conjecture that the optimal scheme is one stage scheme, and a conjecture by Cramer and Bagh (2011) about Shannon minimum and maximum schemes when parent distribution is reflected power. Guidelines for designing an optimal censoring plane are reported together with theoretical and numerical results and illustrations.

Keywords: Progressive type II censoring scheme, Entropy measure, Equivalent schemes, Optimal censoring scheme, Reflected power distribution.

AMS Subject Classification 94A17, 94A20, 62N01, 62N99

*This study is conducted in Al Balqa Applied University during the sabbatical year of the author from The University of Jordan.

1. Introduction

Researchers conduct experiments to collect information about a phenomenon of interest. However, there are several statistical procedures/schemes that lead to loss of information. For example a sample has less information than the underlying probability model, a statistic has less information than the sample and censored sample has less information than the complete sample. The question is how to measure this loss in information and how to reduce it. In this paper, we are concerned with defining a suitable measure of information (objective function) that can select optimal progressive type II censoring schemes and apply it in designing censoring schemes from some parent lifetime distributions.

In a progressive type II censoring scheme, researchers put n units under test, decide to observe only $1 < m \leq n$ failure lifetimes and select a pre-specified censoring scheme $R = (R_1, \dots, R_m)$ that represents the number of units that will be removed, respectively, at the m observed failure times. The class of all possible censoring schemes is

$$CS(m, n) = \{(R_1, \dots, R_m) : R_j \in \{0, 1, \dots, n - m\}, \sum_{j=1}^m R_j = n - m\}.$$

Assume that the life times of these units are independent and identically distributed as a continuous random variable X whose *pdf* and *cdf* are f and F respectively. Let Ω denote the parameter space and $\theta \in \Omega$ be the parameter of the parent distribution of X .

Let $X_{j:m:n}$ denote the j^{th} order statistic in a sequence of m observations under type II progressive censoring scheme. Then for $R \in CS(m, n)$, the pdf of $\vec{X} = (X_{1:m:n}, \dots, X_{m:m:n})$ at the observed sample $\vec{x} = (x_{1:m:n}, \dots, x_{m:m:n})$ is given by

$$f_R(\vec{x}; \theta) = c_R \prod_{j=1}^m f(x_{j:m:n}, \theta) (1 - F(x_{j:m:n}; \theta))^{R_j}, \quad x_{1:m:n} < \dots < x_{m:m:n},$$

where $c_R = \prod_{i=1}^m r_i$ and $r_i = m - i + 1 + \sum_{j=i}^m R_j = n - i + 1 - \sum_{j=1}^{i-1} R_j$.

Let f_C denote pdf of complete sample of order statistics $(X_{1:n}, \dots, X_{n:n})$ when $m = n$ and $C = (R_1, \dots, R_m) = (0, \dots, 0)$. For $1 \leq k \leq m$, $R^{(k)} = (R_1, \dots, R_m)$ denotes a one-stage censoring scheme when $R_i = \begin{cases} n - m & \text{if } i = k \\ 0 & \text{otherwise} \end{cases}$ and its corresponding r_j 's are denote by

$$r_j^{(k)} = \begin{cases} n - j + 1 & \text{if } 1 \leq j \leq k \\ m - j + 1 & \text{if } k + 1 \leq j \leq m \end{cases}$$

The stopping rule of a progressive type II censoring experiment is $T = X_{m:m:n}$. Moreover, the total time on test (Cumulated time of each unit in the experiment) is

$$T_{m:n} = \sum_{j=1}^m r_j (X_{j:m:n} - X_{j-1:m:n}), \text{ with } X_{0:m:n} = 0 \text{ (see e.g. Cramer (2002)).}$$

The problem of selecting an optimal progressive type II censoring scheme has been considered by several researchers.

Burkschat (2008) used ψ -optimality criterion and a partial order relation on $CS(m, n)$ to show that $R^{(m)} \leq R \leq R^{(1)}; \forall R \in CS(m, n)$. Using this result he showed that $R^{(m)}$ minimizes $\psi(R) = E(X_{m:m:n})$. Moreover, if the parent distribution is DFR (decreasing failure rate) then $R^{(m)}$ minimizes both $\psi(R) = E(T_{m:n})$, and $\psi(R) = Var(X_{m:m:n})$. An extra condition is needed for $R^{(m)}$ to minimize $\psi(R) = Var(T_{m:n})$ when distribution is DFR. He also showed that, if the parent distribution is IFR (increasing failure rate), then $R^{(1)}$ minimizes $R = E(T_{m:n})$.

Soliman (2005), obtained MLE and Bayes estimators of reliability, hazard function and parameters of the Burr-XII model, using three loss functions. Based on simulation study, he observed that for fixed m and n , $R^{(1)}$ seems to provide the smallest variance for the estimates of reliability and hazard functions.

The BLUE's of parameters when the underlying distribution is a member of the family

$$\mathfrak{F} = \left\{ F\left(\frac{\# - \mu}{\sigma}\right) : F(x, q) = \begin{cases} 1 - (1 - \text{sign}(q)x)^{1/q}; & q \neq 0 \\ 1 - e^{-x} & ; q = 0 \end{cases} ; x \geq 0, 1 - \text{sign}(q)x > 0, \mu \in \mathbb{R}, \sigma > 0 \right\}$$

had been obtained by several researches. Based on numerical approximation to covariance matrix and numerical search procedures among all possible schemes, it is conjectured that "One-stage censoring scheme is the optimal one". Most frequently $R^{(1)}$ and $R^{(m)}$ are optimal, see e.g. Balakrishnan and Aggarawala (2000), Ng et al. (2002,

2004), Burkschat et al. (2006, 2007), and Balakrishnan et al. (2007). For the same family, Burkschat (2008) considered best linear equivariant estimation and Löwner ordering on mean squared error matrices. Burkschat et al. (2007) also used maximum and minimum eigenvalues of these matrices. In both cases, he concluded that the optimal scheme is either $R^{(1)}$ or $R^{(m)}$ based on the value of q .

For location or scale family, Balakrishnan et al. (2008) suggested the optimality criterion $t(x; \theta) = \left(\frac{\partial}{\partial \theta} \log(\lambda(x; \theta))\right)^2$ where $\lambda(.,.)$ is the failure rate function. They showed that $R^{(1)}$ and $R^{(m)}$ are optimal when $t(., \theta)$ is increasing and decreasing respectively. Moreover, based on numerical study they conjectured that the optimal scheme is $R^{(k)}$ for some k .

Oja (1981) had shown that $R^{(m)}$ minimizes marginal Shannon entropy of duration of experiment, i.e. $H(X_{m:m:n}) = -E(\log f_{X_{m:m:n}}(X))$ if F has DFR.

Based on numerical searching procedure, Awad (2013) showed that the optimal scheme that maximizes the informational efficiency with respect to each of ten \mathbf{A} -sup-entropy measures is $R^{(1)}$ when the underlying distribution is a (shape parameter) Pareto distribution. Same result is obtained theoretically by Haj Ahmad and Awad (2009b) when the underlying distribution is scale-shape Pareto distribution.

Cramer and Bagh (2011) showed that the best approximation to the distribution of an *iid* sample from any distribution is given by the first step censoring $R^{(1)}$ since it minimizes Kullback-Liebler, (KL), divergence measure, Whereas the worst plan results from Type-II right censoring $R^{(m)}$ since it maximized KL from that distribution.

Hofmann et al. (2005), introduced an asymptotic progressive censoring model, and found optimal censoring schemes for location-scale families based on the determinant of the covariance matrix of the asymptotic best linear unbiased estimators. The procedure is illustrated numerically when the parent distributions are Weibull and normal. It is observed that in many situations the obtained optimal schemes significantly improve upon regular Type-II right censoring.

Pradhan and Kundu (2009) used ME algorithm to obtain MLE estimators of the parameters of generalized exponential distribution together with its p^{th} quantile. They applied the missing information principle based on Fisher information matrix. Moreover, they applied the p^{th} quantile information measure. They said that “Till date, we do not have any efficient algorithm to find the optimal censoring scheme in this case.” Then they proposed sub-optimal censoring scheme. Through a simulation study it is observed that the sub-optimal scheme need not be a one-step scheme.

This paper aims at

- a) Suggesting an optimality criterion, based on entropy-information measure, for selecting an optimal progressive type II scheme.

- b) Investigating the suitability of a collection of entropy-information measures as optimality functions based on the suggested criterion.
- c) Exploring the effect of the parameter of the parent distribution on the optimal scheme.
- d) Exploring the uniqueness of optimal censoring scheme together with being one stage scheme.
- e) Exploring the conjecture of Cramer and Bagh (2011) about minimum and maximum Shannon entropy censoring schemes when the underlying distribution is reflected power.
- f) Commenting on the Burkschat (2008) criterion which is based on a partial order relation on the space of all censoring scheme $CS(m, n)$.

Optimality criteria aim at minimizing cost of experimentation, duration of experimentation and/or total time, variability in estimators, or maximizing amount of available information in sampling scheme. In this paper, we are concerned only with objective functions that are based on entropy-information measures.

Let $f_{R;\theta}$ and $f_{C;\theta}$ be joint pdf 's of a progressive type II censored sample based on a scheme R , and a complete sample scheme C . We classify an entropy-information measure I^* , with respect to a given censoring scheme $R \in CS(m, n)$, and a given value of $\theta \in \Omega$ as follows.

- a) If $0 \leq I^*(f_{R;\theta}) < I^*(f_{C;\theta})$, then I^* is called of max-type, (available information - type) with respect to given (R, θ) ,
- b) If $I^*(f_{R;\theta}) > I^*(f_{C;\theta}) \geq 0$, then I^* is called of min-type, (missing information - type) with respect to given (R, θ) ,
- c) If $0 \leq I^*(f_{R;\theta}) = I^*(f_{C;\theta})$, then I^* is called of equivalent-type with respect given (R, θ) .

This classification motivates the following definition of optimality criteria.

Definition 1

- a) For a given θ , the efficiency of a scheme $R \in CS(m, n)$ with respect to I^* is

$$eff_{I^*}(R; \theta) = \min\left\{\frac{I^*(f_{R;\theta})}{I^*(f_{C;\theta})}, \frac{I^*(f_{C;\theta})}{I^*(f_{R;\theta})}\right\}.$$

- b) For a given θ , a censoring scheme $R \in CS(m, n)$ is called I^* - optimal within $CS(m, n)$ if $eff_{I^*}(R; \theta) = \max_{S \in CS(m, n)} \{eff_{I^*}(S; \theta)\}$.
- c) For a given θ , let R_O denote the I^* - optimal scheme and f_O denote the joint pdf of the scheme R_O . Then the relative efficiency of a scheme R with respect to the scheme R_O and the measure I^* is

$$ref_{I^*}(R; \theta) = \min\left\{\frac{I^*(f_{R;\theta})}{I^*(f_{O;\theta})}, \frac{I^*(f_{O;\theta})}{I^*(f_{R;\theta})}\right\}$$

It is clear that $0 < eff_{I^*}(R; \theta) < 1$ and $0 < ref_{I^*}(R; \theta) < 1$. A measure I^* is considered suitable for selecting optimal scheme if it is a function of R . Moreover, a scheme R may be classified as

- a) Low efficient with respect to θ if $eff_{I^*}(R; \theta) < \frac{m}{n}$.
- b) Almost optimal with respect to θ if $ref_{I^*}(R; \theta) \geq 0.98$ and $eff_{I^*}(R; \theta) \geq \frac{m}{n}$.

The following information measures will be used in the sequel

- a) Fisher information (1925): $FI(\vec{X}; \theta) = E\left(\frac{\partial}{\partial \theta} \log(f_R(\vec{X}; \theta))\right)^2$
- b) Kullback-Leibler divergence measure (1951): $KL(f_R; f_S) = E_{f_R}[\log\left(\frac{f_R(\vec{X}; \theta)}{f_S(\vec{X}; \theta)}\right)]$
- c) Shannon entropy (1948): $H(f_R) = -E[\log(f_R(\vec{X}; \theta))]$.
- d) Awad sup-entropy (1987): $AH(f_R) = H(f_R) + \log(s)$, $s = s_R = \sup_{\vec{x}}\{f_R(\vec{x}; \theta)\}$

The paper is organized as follows. Section 2 introduces some preliminary facts about reflected power distribution and generalized order statistics that are used in the paper. Section 3 investigates suitability of Fisher information, Kullback-Leibler divergence, Shannon entropy and Awad sup-modification of Shannon entropy as objective functions to select an optimal progressive type-II censoring scheme when the underlying distribution is reflected power. It also deals with properties of both Shannon entropy and its Awad modification of progressive censoring schemes and provides a partial proof of Cramer and Bagh (2011) conjecture. Moreover, it introduces Shannon maximum and minimum types together with Shannon equivalence of schemes at a given value of θ . Furthermore it deals with properties of Awad-sup entropy together with equivalence of schemes on the parameter space. Section 4 provides results of graphical and numerical computations and comparisons between schemes based on Shannon and Awad-sup entropy measures when parent distribution is reflected power. Finally, Section 5 provides a discussion of obtained results and conclusions of this study.

2. Preliminaries

Let us put n iid units under test that have a reflected power distribution, $(Beta(1, \theta))$, with parameter θ of the probability density function $f(x, \theta) = \theta(1-x)^{\theta-1}$; $0 < x < 1$, $0 < \theta$, cumulative distribution function $F(x, \theta) = 1 - (1-x)^\theta$, failure rate $h(x; \theta) = \frac{f(x; \theta)}{1-F(x; \theta)} = \frac{\theta}{1-x}$ i.e. X has IFR, and quantile function $F^{-1}(x, \theta) = 1 - (1-x)^{1/\theta}$. The joint pdf of the progressive censoring scheme $R = (R_1, \dots, R_m) \in CS(m, n)$ is

$$f_R(\vec{x}; \theta) = c_R \theta^m \prod_{j=1}^m (1 - x_{j:m:n})^{\theta(1+R_j)-1}; \quad 0 < x_{1:m:n} < \dots < x_{m:m:n} < 1 \quad (1)$$

This density is bounded if $\theta > \max\{\frac{1}{1+R_j}; j = 1, \dots, m\}$. Under this condition

$$s_R = s = \max_{\vec{x}} f_R(\vec{x}; \theta) = c_R \theta^m. \quad (2)$$

For the purpose of the evaluation of the above defined information measures, notice that progressive type-II order statistics are special cases of generalized ordered statistics (see e.g. Kamps (1995)). The following lemma is used in the sequel.

Lemma 1 (Kamps 1995)

There exist U_1, \dots, U_m i.i.d. uniform random variables on the interval $[0,1]$ such that $X_{j:m:n} \equiv F^{-1} \left(1 - \prod_{i=1}^j U_i^{\frac{1}{r_i}} \right), j = 1, \dots, m$ in distribution.

Corollary 1: For the progressive type II sample with scheme $R \in CS(m, n)$ from reflected power distribution with parameter θ , there exist U_1, \dots, U_m i.i.d. uniform random variables on the interval $[0,1]$ such

$$a) \quad X_{j:m:n} \equiv 1 - \prod_{i=1}^j U_i^{\frac{1}{\theta r_i}} \text{ in distribution,} \quad (3)$$

$$b) \quad f_R(\vec{X}; \theta) \equiv c_R \theta^m \prod_{i=1}^m U_i^{1 - \frac{(m-i+1)}{\theta r_i}} \text{ in distribution} \quad (4)$$

$$c) \quad f_{R,S}(\vec{X}; \theta) = \frac{f_R(\vec{X}; \theta)}{f_S(\vec{X}; \theta)} \equiv_R \frac{c_R}{c_S} \prod_{i=1}^m U_i^{(1 - \frac{s_i}{r_i})} \text{ in distribution} \quad (5)$$

Proof:

a) Using Lemma 1 and quantile function of distribution it follows that

$$X_{j:m:n} = F^{-1} \left(1 - \prod_{i=1}^j U_i^{\frac{1}{r_i}} \right) = 1 - \prod_{i=1}^j U_i^{\frac{1}{\theta r_i}},$$

b) Put (3) in (1) to get

$$f_R(\vec{X}; \theta) \equiv c_R \theta^m \prod_{j=1}^m \left(\prod_{i=j}^m U_i^{\frac{\theta(1+R_j)-1}{\theta r_i}} \right),$$

$$\text{Since } \prod_{j=1}^m \prod_{i=1}^j a_i^{b_{i,j}} = \prod_{i=1}^m a_i^{\sum_{j=i}^m b_{i,j}}$$

$$f_R(\vec{X}; \theta) \equiv c_R \theta^m \prod_{i=1}^m U_i^{\sum_{j=i}^m \frac{\theta(1+R_j)-1}{\theta r_i}}$$

$$f_R(\vec{X}; \theta) \equiv c_R \theta^m \prod_{i=1}^m U_i^{1 - \frac{(m-i+1)}{\theta r_i}}$$

$$c) \quad f_{R,S}(\vec{X}; \theta) = \frac{f_R(\vec{X}; \theta)}{f_S(\vec{X}; \theta)} = \frac{c_R}{c_S} \prod_{j=1}^m (1 - X_j)^{\theta(R_j - S_j)}$$

$$f_{R,S}(\vec{X}; \theta) \equiv_R \frac{c_R}{c_S} \prod_{j=1}^m \left(\prod_{i=1}^j U_i^{\frac{1}{\theta r_i}} \right)^{\theta(R_j - S_j)} = \frac{c_R}{c_S} \prod_{i=1}^m U_i^{(1 - \frac{s_i}{r_i})}$$

Corollary 2:

Consider the progressive type II sample with scheme $R \in CS(m, n)$ from reflected power distribution with parameter θ . Then

$$E \left(\log \left(f_R(\vec{X}; \theta) \right) \right) = \log(c_R) + m \log(\theta) - m + \frac{m}{\theta} + \frac{1}{\theta} \sum_{i=2}^m \frac{(m-i+1)}{r_i}, \quad (6)$$

$$E_R[\log(f_{R,S}(\vec{X}; \theta))] = \log\left(\frac{c_R}{c_S}\right) - \sum_{i=2}^m \left(1 - \frac{s_i}{r_i}\right) = \sum_{i=2}^m \left\{ \frac{s_i}{r_i} - \log\left(\frac{s_i}{r_i}\right) - 1 \right\} \quad (7)$$

Proof

All required forms follow by using corollary 1, the independence of the involved uniform random variables, and the facts that $E(-\log(U)) = 1$ and $E(U^\eta) = \frac{1}{\eta+1}$; $\eta > -1$

3. Information Measures of Progressive censored scheme

The first condition on an information (objective) function to be suitable for selecting an optimal censoring scheme is being a function of the censoring scheme R . This section aims at deciding which of the above defined information measures satisfy this first condition.

3.1 Fisher Information

Let us investigate the suitability of Fisher information through the following

Lemma 2:

For any progressive type-II censoring scheme $R \in CS(m, n)$, when the underlying distribution is reflected power with parameter θ , the Fisher information $I_1(f_R; \theta) = \frac{m}{\theta^2}$

Proof:

Take the logarithm of (1), to get

$$f_R^* = \log(c_R) + m \log(\theta) + \theta \sum_{j=1}^m (1 + R_j) \log(1 - x_{j:m:n}) - \sum_{j=1}^m \log(1 - x_{j:m:n})$$

$$\text{Hence } \frac{\partial^2 f_R^*}{\partial \theta^2} = -\frac{m}{\theta^2}, \quad I_1(R; \theta) = \frac{m}{\theta^2} \quad \text{and} \quad I_1(f_C; \theta) = \frac{n}{\theta^2}$$

Therefore Fisher information measure is of maximum-type but it cannot pick an optimal scheme since it is free of the censoring scheme R . Moreover, the efficiency of any scheme is

$$m/n.$$

3.2 Divergence Measures

The Kullback-Leibler divergence measure between f_R and f_S is given (7). So, it is a function of R and S through r_i and s_i for $i = 2, \dots, m$. This measure looks like the

well-known entropy loss function, So, $KL(S; R)$ may be interpreted as the total loss of using S as an alternative to R . The symmetric form of this measure is

$$J(R; S) = \frac{1}{2} (KL(R; S) + KL(S; R)) = 1 - m + \frac{1}{2} \sum_{i=2}^m \left(\frac{s_i}{r_i} + \frac{r_i}{s_i} \right)$$

The smaller this value is the more closely are the two schemes. So, it may be used to find alternative schemes of an existing traditionally used one. This type of problem will be fully investigated in another paper.

3.3 Entropy Measures

In this section we investigate optimality and efficiency of progressive type-II schemes with respect to two entropy measures.

3.3.1 Shannon Entropy

Corollary 3:

Consider the progressive type II sample with scheme $R \in CS(m, n)$ from reflected power distribution with parameter θ . Then

$$a) \quad H(f_R) = -\log(c_R) - m \log(\theta) + m - \frac{m}{\theta n} - \frac{1}{\theta} \sum_{i=2}^m \frac{(m-i+1)}{r_i} \quad (8)$$

$$b) \quad H(f_C) = -\log(n!) - n \log(\theta) + n - \frac{n}{\theta}, \quad (9)$$

$$c) \quad \text{If } \theta > \max\left\{\frac{1}{1+R_j}; j = 1, \dots, m\right\} \text{ then } AH(f_R) = m - \frac{1}{\theta} \frac{m}{n} - \frac{1}{\theta} \sum_{i=2}^m \frac{m-i+1}{r_i} \quad (10)$$

$$d) \quad \text{If } \theta > 1 \text{ then for complete sample } AH(f_C) = n(1 - \frac{1}{\theta}), \quad (11)$$

Proof:

Use the definitions of H and AH entropy measures together with (6) and (2) to get (8) and (10). In (8) and (10), put $m = n$, and $R_j = 0$; for $j = 1, \dots, m$, i.e. $r_i = n - i + 1$ to get (9) and (11).

$H(f_R)$ and $AH(f_R)$ are functions of R through $r'_i s$, $i = 2, \dots, m$. So, it is worth investigating them as information functions that may be used to select optimal scheme. This is done in the rest of the paper.

It should be mentioned that Cramer and Bagh (2011) treated reflected power distribution using Shannon entropy. However, the treatment and the results in this paper are completely different from those of Cramer and Bagh.

Theorem 1:

Let $H(f_R)$, f_R , and f_C denote Shannon entropy measure, joint *pdf* of progressive type II censoring scheme $R \in CS(m, n)$, and joint *pdf* of complete first n order statistics when the parent distribution is reflected power ($Beta(1, \theta)$). Then

1. $\forall R \in CS(m, n); 0 > H(f_R; \theta) > H(f_C; \theta).$ (12)
2. The Shannon efficiency of R is $eff_H(R) = \frac{H(f_R; \theta)}{H(f_C; \theta)}$ which takes value between zero and one. Moreover, $\lim_{\theta \rightarrow \infty} eff(H(f_R; \theta)) = \frac{m}{n}$ and $\lim_{\theta \rightarrow 0} eff(H(f_R; \theta)) = \frac{1}{n} \sum_{i=1}^m \frac{m-i+1}{r_i} < \frac{m}{n}.$
3. There is no scheme R that is H-equivalent to the complete scheme C . However, if the true value of $\theta = \frac{n - \sum_{i=1}^m \frac{(m-i+1)}{r_i}}{n-m} = \theta_{0,R}, (say),$ then R will be the most efficient scheme within $CS(m, n).$

Proof:

Since θ plays an important role in the proof, set

$$g(\theta) = H(f_R; \theta) = H(f_R) = -\log(c_R) - m \log(\theta) + m - \frac{1}{\theta} \sum_{i=1}^m \frac{(m-i+1)}{r_i}$$

Then $\frac{\partial g}{\partial \theta} = \frac{m}{\theta^2} (\frac{1}{m} \sum_{i=1}^m \frac{m-i+1}{r_i} - \theta)$ which equals zero at $\theta = \frac{1}{m} \sum_{i=1}^m \frac{m-i+1}{r_i} = \theta_{1,R}, (say).$ Note that $0 < \theta_{1,R} < 1.$ Moreover, $g(\theta)$ is increasing when $\theta < \theta_{1,R},$ decreasing when $\theta > \theta_{1,R},$ and the maximum value of g is

$$g(\theta_{1,R}) = -\log(c_R) - m \log(\theta_{1,R}) = -\log(c_R \theta_{1,R}^m)$$

By arithmetic-geometric mean inequality applied to the sequence

$$b_i = \frac{(m-i+1)}{r_i}, \text{ we get } \theta_{1,R} = \frac{1}{m} \sum_{i=1}^m \frac{(m-i+1)}{r_i} \geq (\prod_{i=1}^m \frac{(m-i+1)}{r_i})^{\frac{1}{m}},$$

$$(\prod_{i=1}^m \frac{(m-i+1)}{r_i})^{\frac{1}{m}} = (\frac{m!}{\prod_{i=1}^m r_i})^{\frac{1}{m}} \geq \frac{1}{(\prod_{i=1}^m r_i)^{\frac{1}{m}}} = \frac{1}{(c_R)^{1/m}}$$

These two inequalities imply that $\theta_{1,R} c_R^{1/m} \geq 1,$ i.e. $\theta_{1,R}^m c_R \geq 1$ and hence $g(\theta_{1,R})$ the maximum value of $H(f_R, \theta)$ is negative. Moreover, $\lim_{\theta \rightarrow 0} g(\theta) = \lim_{\theta \rightarrow \infty} g(\theta) = -\infty.$ Hence $\forall R \in CS(m, n), H(f_R, \theta) < 0.$

Rearrange terms of $H(f_C)$ in (9), and use the fact that $\log(t) \leq t - 1$ to get

$$H(f_C) = -\log(n!) + n \left\{ \log\left(\frac{1}{\theta}\right) - \frac{1}{\theta} + 1 \right\} < 0.$$

Now, set $\Delta H(\theta) = H(f_R) - H(f_C),$ use (8) and (9) and then rearrange terms to get

$$\Delta H(\theta) = \log\left(\frac{n!}{c_R}\right) - (n-m) \left[\log\left(\frac{1}{\theta}\right) + 1 - \frac{1}{\theta} \left\{ \frac{n - \sum_{i=1}^m \frac{(m-i+1)}{r_i}}{n-m} \right\} \right]$$

Note that $\frac{n!}{c_R} > 1$, $\sum_{i=1}^m \frac{(m-i+1)}{r_i} < m$ and use $\log\left(\frac{1}{\theta}\right) \leq \frac{1}{\theta} - 1$ with equality iff $\frac{1}{\theta} = 1$ to get $\Delta H(\theta) > 0$. Hence (12) follows.

The second part follows by dividing the inequalities in the (12) by $H(f_C)$ and using the L'Hospital rule to calculate the limits.

Finally, direct differentiation shows that $\Delta H(\theta)$ has a minimum value at $\theta = \theta_{0,R}$.

The minimum value is $\Delta H(\theta_{0,R}) = \log\left(\frac{n!}{c_R}\right) + (n-m) \log\left(\frac{n - \sum_{i=1}^m \frac{(m-i+1)}{r_i}}{n-m}\right)$.

Since $\Delta H(\theta_{0,R}) > 0$, there is no scheme R that is Shannon equivalent to the complete scheme C . However, if the true value of $\theta = \theta_{0,R}$, then the scheme R will be the most efficient scheme within $CS(m, n)$ since it is closest to C .

Theorem 2:

Let R and S be two schemes from $CS(m, n)$. Use standard notations of r_i , c_R and the corresponding s_i and c_S . Set $H\Delta(R, S) = H(f_R; \theta) - H(f_S; \theta)$. Then

$$(a) \quad H(f_R; \theta) = H(f_S; \theta) \quad \text{iff} \quad \theta = \theta_{0,R,S} = \frac{\sum_{j=2}^m (m-j+1) \left(\frac{1}{s_j} - \frac{1}{r_j}\right)}{\sum_{j=2}^m \log\left(\frac{r_j}{s_j}\right)}; \quad \text{provided that}$$

$\sum_{j=2}^m \log\left(\frac{r_j}{s_j}\right) \neq 0$, otherwise there is no θ such value $H(f_R) = H(f_S)$.

(b) If $c_R > c_S$ and $\sum_{j=2}^m \log\left(\frac{s_j}{r_j}\right) \neq 0$ then

(i) $H(R; \theta) < H(S; \theta)$ when $\theta_{0,R,S} < \theta$

(ii) $H(R; \theta) > H(S; \theta)$ when $\theta_{0,R,S} > \theta$

Proof

(a) Use (8) and the condition $\sum_{j=2}^m \log\left(\frac{r_j}{s_j}\right) \neq 0$ to get

$$H(f_R; \theta) = H(f_S; \theta) \Leftrightarrow \log\left(\frac{c_R}{c_S}\right) + \frac{1}{\theta} \sum_{j=1}^m (m-j+1) \left(\frac{1}{r_j} - \frac{1}{s_j}\right) = 0,$$

Since, $r_1 = s_1 = n$, and $c_R = \prod_{j=1}^m r_j$, part (a) follows.

(Note that $0 < \theta_{0,R,S}$ since numerator and denominator have same sign.)

(b) Note that $H\Delta(R, S) = \frac{1}{\theta} \sum_{j=2}^m (m-j+1) \left(\frac{1}{s_j} - \frac{1}{r_j}\right) + \sum_{j=2}^m \log\left(\frac{s_j}{r_j}\right)$ and the condition $c_R > c_S$ is equivalent to $r_i > s_i$; $i = 2, \dots, m$. Hence

$$H\Delta(R, S) = \left\{1 - \frac{\theta_{0,R,S}}{\theta}\right\} \sum_{j=2}^m \log\left(\frac{s_j}{r_j}\right)$$

Since $\sum_{j=2}^m \log\left(\frac{s_j}{r_j}\right) < 0$, the required results follows.

Remark 1:

Mathematical10 code is used to illustrate Theorem 2. From the obtained output it is observed that

- 1) When $\theta=0.420635$, $H(\{2,0,2\}, \theta) \equiv H(\{0,3,1\}, \theta)$ even though $\sum_{j=2}^m \log\left(\frac{r_j}{s_j}\right) = 0$. Therefore the condition on equivalence of the two schemes (Theorem 2), that $\sum_{j=2}^m \log\left(\frac{r_j}{s_j}\right) \neq 0$ is sufficient but not necessary.
- 2) There may be more than two schemes that are Shannon equivalent at the same value of θ . For example, at $\theta = 0.214918$, $\{2,0,5\} \equiv \{2,1,4\} \equiv R^{(m)}$, and at $\theta = 0.606826$, $\{1,2,2\} \equiv \{1,3,1\} \equiv R^{(1)}$. This means that even if $R^{(1)}$ or $R^{(m)}$ are minimum or maximum Shannon plans, then they may not be unique.

Corollary 4:

- (a) $\forall k < l, \quad H(f_{R^{(k)}}; \theta) = H(f_{R^{(l)}}; \theta) \text{ iff } \theta = \theta_1 = \frac{(n-m) \sum_{j=k+1}^l \frac{1}{n-j+1}}{\sum_{j=k+1}^l \log\left(1 + \frac{n-m}{m-j+1}\right)}$
- (b) $H(f_{R^{(1)}}; \theta) = H(f_{R^{(m)}}; \theta) \text{ iff } \theta = \theta_2 = \frac{(n-m) \sum_{j=2}^m \frac{1}{n-j+1}}{\sum_{j=2}^m \log\left(1 + \frac{n-m}{m-j+1}\right)}$

Proof:

Set $R = R^{(k)}$ and $S = R^{(l)}$, then the condition $c_R > c_S$ holds. So, apply part (b) of Theorem 2 to R and S , then rewrite $\frac{m-j+1}{n-j+1} = 1 - \frac{n-m}{n-j+1}$, $\frac{n-j+1}{m-j+1} = 1 + \frac{n-m}{m-j+1}$ and $\log\left(\frac{m-j+1}{n-j+1}\right) = -\log\left(1 + \frac{n-m}{m-j+1}\right)$, to get (a). Part (b) follows from (a) when $k = 1$ and $l = m$.

Corollary 5:

If $R = R^{(m)}$ & $S = R^{(1)}$ then

- a) $\theta_{0,R,S} = \theta^* < 1$ (θ^* appeared in conjecture of Cramer and Bagh (2011))
- b) $\sum_{j=2}^m \log\left(\frac{r_j}{s_j}\right) \neq 0$
- c) $c_R > c_S$
- d) $H(R^{(m)}; \theta) < H(R^{(1)}; \theta)$ when $\theta^* < \theta$
- e) $H(R^{(m)}; \theta) > H(R^{(1)}; \theta)$ when $1 > \theta^* > \theta$

So, this provides a proof of the second part of Cramer and Bagh (2011) conjecture, i.e.

“In particular, $R^{(1)}$ yields a smaller entropy than $R^{(m)}$ for $\theta \leq \theta^* = \frac{(n-m) \sum_{j=2}^m \frac{1}{n-j+1}}{\sum_{j=2}^m \log\left(1 + \frac{n-m}{m-j+1}\right)} \in (0,1)$; $n > m \geq 2$. For $\theta \geq \theta^*$; $R^{(m)}$ leads to a smaller entropy than $R^{(1)}$.”

Remark 2:

- 1) Part (b) of Theorem 2 shows that if $c_R > c_S$ then it is not necessarily true that S is more informative than R as it may be concluded from the partial order relation given by Burkschat (2008).
- 2) For $\theta < 1$, the curves of $H(f_{R^{(1)}}, \theta)$ & $H(f_{R^{(m)}}, \theta)$ concaves downward and intersect at $\theta_2 = \theta^*$ and the entropy curve of any other scheme $H(f_R, \theta)$ is also concaves downward and will intersect each of $H(f_{R^{(1)}}, \theta)$ & $H(f_{R^{(m)}}, \theta)$ at $\theta_{0,R,R^{(1)}}$ and $\theta_{0,R,R^{(m)}}$ respectively. There is always a scheme R such that $H(R; \theta^*)$ is either less than or greater than $H(f_{R^{(1)}}, \theta^*) = H(f_{R^{(m)}}, \theta^*)$. This remark supports the second part of Cramer and Bagh (2011) conjecture.

3.3.2. Awad sup-entropy AH

Theorem 3

Let $AH(f_R; \theta)$, f_R , and f_C denote Awad modification of Shannon entropy measure, joint *pdf* of progressive type II censoring scheme $R \in CS(m, n)$, and joint *pdf* of complete first n order statistics when the parent distribution is reflected power ($Beta(1, \theta)$). Set $B = \max\{\frac{1}{1+R_j}; j = 1, \dots, m\}$, and $\theta_{0,R} = \frac{n - \sum_{i=1}^m \frac{(m-i+1)}{r_i}}{n-m}$. Assume that > 1 . Then $\forall R \in CS(m, n)$;

- a) A scheme R is A-equivalent to the complete scheme at $\theta = \theta_{0,R}$
- b) $0 < AH(f_R; \theta) \leq AH(f_C; \theta)$ when $\theta > \theta_{0,R}$, but
 $AH(f_R; \theta) > AH(f_C; \theta) > 0$ when $1 < \theta < \theta_{0,R}$.

Proof:

It is obvious that $AH(f_C; \theta) > 0$ and $\forall R \in CS(m, n)$; $AH(f_R; \theta) > 0$ since each is the negative of the expected value of logarithm of a plausibility function which is bounded between zero and one.

Set $\Delta AH(\theta) = AH(f_R; \theta) - AH(f_C; \theta)$. The forms of $AH(f_R; \theta)$ and $AH(f_C; \theta)$ in (10) and (11), imply that

$$\Delta AH(\theta) = -(n-m) \left[1 - \frac{1}{\theta} \left\{ \frac{n - \sum_{i=1}^m \frac{(m-i+1)}{r_i}}{n-m} \right\} \right]; \theta > 1.$$

Hence $\Delta AH(\theta) = 0$ when $\theta = \frac{n - \sum_{i=1}^m \frac{(m-i+1)}{r_i}}{n-m} = \theta_{0,R}$. Hence part (a) is proved.

So, $\Delta AH(\theta) = (n - m) \left[\frac{\theta_{0,R}}{\theta} - 1 \right]$, and hence part (b) follows since $\theta > 1$ and $\theta_{0,R} > 1$.

Remark 3:

- 1) Theorem 3 implies that $\forall \theta, \forall R \in CS(m, n)$ Awad sup- entropy is of maximum type when $\theta > \theta_{0,R}$, of minimum type when $1 < \theta < \theta_{0,R}$, and equivalent type when $\theta = \theta_{0,R}$.
- 2) A scheme R is equivalent to the complete scheme C at $\theta = \theta_{0,R}$. It is interesting to note that this point is also the point at which R is the most Shannon efficient scheme within the class $CS(m, n)$.
- 3) $AH(f_C; \theta)$ is defined when $\theta > 1$ but $AH(f_C; \theta)$ is defined when $\theta \in (B, \infty)$ where $B = \max\{1/(1 + R_j); j = 1, \dots, m\} < 1$.
- 4) $\lim_{\theta \rightarrow \infty} AH(f_R; \theta) = m$, and $\lim_{\theta \rightarrow \infty} AH(f_C; \theta) = n$
- 5) $\lim_{\theta \rightarrow 1^+} AH(f_R; \theta) = m - \sum_{i=1}^m \frac{(m-i+1)}{r_i}$, and $\lim_{\theta \rightarrow 1^+} AH(f_C; \theta) = 0$
- 6) Direct differentiation implies that $AH(f_R; \theta)$ and $AH(f_C; \theta)$ are increasing and concave downward functions in θ .

Theorem 4:

1. $\forall \theta > 1$, $AH(f_R; \theta) = AH(f_S; \theta)$ iff $\sum_{i=2}^m (m - i + 1) \left(\frac{1}{r_i} - \frac{1}{s_i} \right) = 0$
2. If $r_j < s_j \quad \forall j = 2, \dots, m$; then $AH(f_R; \theta) < AH(f_S; \theta)$.
3. There are no pairs of single stage schemes that are AH -equivalent.
4. $\forall \theta > 1$, $AH(f_{R^{(1)}}; \theta) < AH(f_{R^{(m)}}; \theta)$

Proof:

Set $\Delta \Delta(R, S) = AH(f_R; \theta) - AH(f_S; \theta)$. Then, use (10) to get

$$\Delta \Delta(R, S) = \frac{1}{\theta} \sum_{j=2}^m (m - j + 1) \left(\frac{1}{s_j} - \frac{1}{r_j} \right). \text{ Hence, first part follows.}$$

The second part follows from the first one and the condition $r_j < s_j \quad \forall j = 2, \dots, m$.

For the third part, by contradiction, assume that there is a $1 \leq k < l \leq m < n$ such that $AH(R^{(k)}; \theta) = AH(R^{(l)}; \theta)$.

$$\text{This implies that } \sum_{i=2}^m (m - i + 1) \left(\frac{1}{r_i^{(l)}} - \frac{1}{r_i^{(k)}} \right) = 0 \quad (13)$$

On the other hand, it is clear that

$$\frac{1}{r_j^{(l)}} - \frac{1}{r_j^{(k)}} = \begin{cases} \frac{1}{n-j+1} - \frac{1}{m-j+1} & \text{if } k+1 \leq j \leq l, \text{ and} \\ 0 & \text{otherwise} \end{cases}$$

$$(m-j+1)\left(\frac{1}{r_j^{(l)}} - \frac{1}{r_j^{(k)}}\right) = \begin{cases} -\frac{n-m}{n-j+1} < 0 & \text{if } k+1 \leq j \leq l \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

It is clear that (13) contradicts (14). So, the required result is obtained.

Finally, for the last part,

$$AH(f_{R^{(1)}}; \theta) = m - \frac{1}{\theta} \frac{m}{n} - \frac{1}{\theta} \sum_{i=2}^m \frac{m-i+1}{m-i+1} = m - \frac{1}{\theta} \frac{m}{n} - \frac{1}{\theta} (m-1), \quad (15)$$

$$AH(f_{R^{(m)}}; \theta) = m - \frac{1}{\theta} \sum_{i=1}^m \frac{m-i+1}{n-i+1} = m - \frac{1}{\theta} \frac{m}{n} - \frac{1}{\theta} \sum_{i=2}^m \frac{m-i+1}{n-i+1} \quad (16)$$

The required result follows since $\sum_{i=2}^m \frac{m-i+1}{n-i+1} < (m-1)$.

Remark 4:

1. It is interesting to note that the condition on equivalence of a pair of censoring schemes is free of the parameter θ . This means that if the condition $\sum_{i=2}^m (m-i+1)\left(\frac{1}{r_i} - \frac{1}{s_i}\right) = 0$ holds then the two schemes R and S are AH -equivalent for all θ .
2. For a given θ , if AH is a measure of maximum type with respect to S then S is more informative than R . Otherwise if it is a measure of minimum type with respect to R then scheme R is more informative than S .
3. Table 2 reports all AH -equivalent pairs of schemes when $n = 5, \dots, 10$ and $m = 2, \dots, n-1$. For these values of n and m , it is interesting to note that there are no pairs of AH equivalent schemes other than those reported in Table 2.

Corollary 6:

$$\forall R \in CS(m, n) \text{ \& } \forall \theta;$$

$$AH(f_{R^{(1)}}; \theta) \leq AH(f_R; \theta) \leq AH(f_{R^{(m)}}; \theta)$$

Proof

Note that $\forall R \in CS(m, n)$, $r_1 = n$, and $\forall i \geq 2$; $m-i+1 \leq r_i \leq n-i+1$. (17)

$$\text{Hence, } \min_{R \in CS(m, n)} \frac{m-i+1}{r_i} = \frac{m-i+1}{n-i+1}; \quad \forall i \geq 1 \quad \text{and} \quad \max_{R \in CS(m, n)} \frac{m-i+1}{r_i} = \begin{cases} \frac{m}{n} & \text{if } i = 1 \\ 1 & \text{if } i \geq 2 \end{cases}$$

Now using the form of $AH(f_R; \theta)$, (15) and (16), it follows that

$$\begin{aligned}\max_{R \in CS(m,n)} AH(f_R; \theta) &= m - \frac{m}{\theta n} - \frac{1}{\theta} \sum_{i=2}^m \frac{m-i+1}{n-i+1} = AH(f_{R^{(m)}}; \theta), \\ \min_{R \in CS(m,n)} AH(f_R; \theta) &= m - \frac{1}{\theta} \frac{m}{n} - \frac{1}{\theta} (m-1) = AH(f_{R^{(1)}}; \theta),\end{aligned}$$

Hence, $\forall R \in CS(m, n)$,

$$AH(f_{R^{(1)}}; \theta) \leq AH(f_R; \theta) \leq AH(f_{R^{(m)}}; \theta).$$

Remark 5:

It is interesting to note that, this result neither implies $R^{(1)}$ nor $R^{(m)}$ is the optimal scheme since the decision depends on the entropy AH being min- or max- type with respect to these schemes and the parameter θ . In our terminology the scheme whose entropy is the closest to the entropy of the complete sample is the optimal one.

5. Numerical Computations

To explore the Cramer and Bagh (2011) conjecture and illustrate the use of the obtained theoretical results we have developed a Mathematica10 code to do all necessary illustrations. Typical part of output of this code is reported in the appendix that includes the following parts;

1. Table 1 reports the values of Shannon entropy in vicinity of θ^* , (that appeared in conjecture of Cramer and Bagh (2011)), i.e. at $\theta^* - \epsilon$, θ^* , $\theta^* + \epsilon$ for all schemes in $CS(3,8)$, where $\epsilon = 0.1$. Note that (θ^* is the same for all schemes in $CS(m, n)$). From Table 1 we observe that
 - a) At θ^* there is only one pair of schemes $\{0, 0, 5\}$ and $\{5, 0, 0\}$ that are Shannon equivalent. These two schemes are minimum Shannon plan, i.e. such a plan is not unique.
 - b) For the scheme $\{0,0,5\}$, $H(\theta^* - \epsilon) > H(\theta^*) > H(\theta^* + \epsilon)$
 - c) For the scheme $\{5,0,0\}$, $H(\theta^* - \epsilon) < H(\theta^*) < H(\theta^* + \epsilon)$. Same holds for each of the following schemes $\{1, 4, 0\}, \{2, 3, 0\}, \{3, 2, 0\}, \{4, 1, 0\}$
 - d) For the scheme $\{3, 1, 1\}$, $H(\theta^* + \epsilon) < H(\theta^*) < H(\theta^* - \epsilon)$. At θ^* , this scheme is the unique maximum Shannon plan. It is also not one stage scheme.
 - e) For the schemes $\{0,5,0\}$ and $\{4,0,1\}$, $H(\theta^* - \epsilon) < H(\theta^*) < H(\theta^* + \epsilon)$
 - f) For all other schemes in $CS(3,8)$ that are not mentioned above, the above inequalities in (a) of scheme $\{0,0,5\}$ hold.
 - g) When θ is close enough to θ^* , the minimum scheme when $\theta < \theta^*$ is $\{5,0,0\}$ while it is $\{0,0,5\}$ when $\theta > \theta^*$. On the other hand, the maximum scheme is $\{2,1,2\}$ when $\theta < \theta^*$, while it is $\{4,0,1\}$ when $\theta > \theta^*$.
 - h) This indicates how the minimum and maximum schemes are very sensitive to the value of the unknown parameter of the parent distribution.

2. Table 2 reports all pairs of schemes (R, S) such that R and S are A -equivalent for $n = 7(1)10$ and m takes all possible values between 2 and $n - 1$.
3. Table 3 reports values of θ and $R \in CS(5,10)$ at which $\Delta HRC = H(R) - H(C)$ is minimum and the H -efficiency of the scheme is greater than $m/n = 0.50$. At the same values of θ , each of those schemes is A -equivalent to the complete scheme C . It is observed that total number of schemes is 126 out of which only 13 schemes can lead to H -efficiency more than $m/n = 0.50$ and the values of θ belongs to the sub-parameter space $[1.5954, 1.64563]$. This suggests that if someone has prior information that the value of θ is in this sub-space, and he is planning to put 10 units under test and to observe only 5 failure times, then he may only take into consideration these 13 schemes.
4. Figure 1a provides typical plots of $H(R; \theta)$, $H(C; \theta)$, $eff_H(R; \theta)$, and m/n , while figure 1b provides typical plots of $AH(R; \theta)$, $AH(C; \theta)$, $eff_A(R; \theta)$, and m/n when $R = \{0,0,1,1,3\}$. It is seen from Figure 1b that the curves of $(R; \theta)$, and $AH(C; \theta)$ intersect at exactly one point. This means that the censoring scheme may be minimum, maximum or equivalent type based on value of θ . From the curve of $eff_A(R; \theta)$, the partition points of parameter space are $\{\frac{1303}{1080}, \frac{583}{360}\}$, i.e. If $\theta < \frac{1303}{1080}$ then the scheme R has low A -efficiency which is less than $\frac{m}{n} = 0.5$, otherwise the efficiency is more than 0.5. The A -efficiency is one at exactly one value of $\theta = \frac{583}{360}$. If θ is to the left of the previous value then the scheme is minimum type and if it is to the right then the scheme is maximum type. Furthermore the A -efficiency starts decreasing for $\theta > \frac{583}{360}$.
5. Figure 1c: provides typical plots of A -entropies and A -efficiencies of $R^{(1)}, R^{(m)}$, and $R = \{1,1,1,0,2\}$. It is observed from these tables and solutions for cut points of the obtained curves that each of $R^{(1)}, R^{(m)}$, and R , has A -efficiency equals one at $\theta = \frac{11}{10}, \frac{23}{15}$, and $\frac{4147}{2520}$ respectively. In addition, among these three schemes, the most informative scheme is $R^{(1)}$ when $\theta \in (1, 1.27561)$, R when $\theta \in (1.27561, 1.58814)$, otherwise it is $R^{(m)}$.

4. Conclusions and Recommendations

In this paper, it is proved that

- a) Any two one-stage censoring schemes R and S from $CS(m, n)$ are not Awad sup-entropy equivalent. However, they are Shannon equivalent at a single value of θ .
- b) Any schemes R and S from $CS(m, n)$ may be Awad equivalent for all values of $\theta > 1$. However, they may be Shannon equivalent at a single value of θ . Moreover, for a given value of θ there may exist more than two schemes that are H -equivalent.
- c) Each scheme R is Awad equivalent to C at the single value

$$\theta_{0,R} = \frac{n - \sum_{i=1}^m \frac{(m-i+1)}{r_i}}{n-m} > 1.$$

There is no scheme R that is Shannon equivalent to the complete scheme C . However, R is the most H-efficient scheme when $\theta = \theta_{0,R}$, in the sense that the minimum distance between the curves of $H(R, \theta)$ and $H(C, \theta)$ occurs at $\theta_{0,R}$.

d) $\forall \theta, \forall R \in CS(m, n); 0 > H(f_R; \theta) > H(f_C; \theta).$

However, $\forall R \in CS(m, n); 0 < AH(f_R; \theta) \leq AH(f_C; \theta)$ when $\theta > \theta_{0,R}$, but $AH(f_R; \theta) > AH(f_C; \theta) > 0$ when $1 < \theta < \theta_{0,R}$.

e) $\forall \theta > 1$ and $\forall R \in CS(m, n); AH(R^{(1)}) \leq AH(R) \leq AH(R^{(m)})$. However, this does not mean that $R^{(m)}$ is the most informative scheme, and it also does not mean that $R^{(1)}$ is the least informative one within the class $CS(m, n)$. To be more specific, let $R \in (CS(m, n) - \{R^{(m)}, R^{(1)}\})$ then it is clear that

a) If $\theta > \theta_{0,R^{(m)}}$, then $R^{(m)}$ is of maximum type and it is the A-optimal scheme within the class $CS(m, n)$.

b) If $\theta < \theta_{0,R^{(1)}}$, then $R^{(1)}$ is of minimum type, and it is the A-optimal scheme within the class $CS(m, n)$.

c) If $\theta_{0,R} < \theta < \theta_{0,R^{(m)}}$, then R is maximum type while $R^{(m)}$ is minimum type, and the A-optimal scheme is either $R^{(m)}$ or an R within the class $(CS(m, n) - \{R^{(m)}, R^{(1)}\})$.

d) If $\theta_{0,R^{(1)}} < \theta < \theta_{0,R}$, then R is minimum type while $R^{(1)}$ is maximum type, and the A-optimal scheme either $R^{(1)}$ or an R within the class $(CS(m, n) - \{R^{(m)}, R^{(1)}\})$.

f) Concerning Shannon entropy it is observed that

1. The value of θ^* that in Cramer-Bagh (2011) conjecture is the point at which $R^{(1)}$ and $R^{(m)}$ are Shannon entropy equivalent.
2. The maximum and minimum Shannon entropy schemes are sensitive to the value of the parameter θ . They need not be unique, neither being $R^{(1)}$ or $R^{(m)}$.

g) The suggested definition of efficiency that is based on classifying schemes as being of maximum, minimum, or equivalent type seems to be more appropriate than that definition that just compares the values of the entropy in the schemes without taking this issue into consideration.

h) At the design stage of a progressive censoring plan one may take into consideration the following remarks.

- 1) If someone observed failure times of m units out of the n units that under test, then it seems reasonable to claim that any scheme with efficiency less than m/n is a scheme with low efficiency. In such case, it is not preferable to use censoring since any inference based on low efficient scheme will not be beneficial.

- 2) A prior information (from past experience or a pilot study on complete samples) about the possible value of the unknown parameter is very helpful in selecting optimal scheme, since optimal schemes are sensitive to the values of θ .
- 3) Since, for a given θ , there may be no A- or H-equivalent scheme, one may find an equivalent or almost efficient scheme for some θ in the vicinity of the prior value of θ and searches for a most efficient scheme.
- 4) The researcher is advised to find all equivalent schemes to the selected optimal one and then to use his judgment based on practical situations and objectives of his experiment to use the most appropriate one.

Finally for a given underlying distribution, one may construct tables that report $n, m, \theta \in \Omega$ and efficiency of $R \in CS(m, n)$ that are greater than m/n . All equivalent schemes should be grouped. Based on such a table, the researcher may use the prior subspace of values of θ to select m, n and R that may lead to efficiency greater than m/n . If there is no scheme with efficiency greater than m/n , it is advised to use censoring.

Acknowledgments

The author is gratefully thankful for The University of Jordan for giving him the opportunity to complete this part of his research project during his sabbatical year. He also thanks Al-Balqa Applied University for allowing him to use its facilities during his sabbatical year from The University of Jordan.

References

1. Abo-Eleneen, Z. A., (2012). A novel approach for optimal schemes in progressive censoring plans. J. Comm. Comput., 9, 426–433.
2. Awad, A.M., (1987). “A Statistical Information Measure”, *Dirasat (Science)*, XIV (12), pp. 7-20.
3. Awad, A. M., and Alwanah, A. (1987). Application of Entropy to a Life-Time Model, IMA J. of Mathematical Control and Information 4, 143-147.
4. Awad, A. M. (2013) Informational Criterion of Censored Data Part I: Progressive Censoring Type II with Application to Pareto Distribution, In: Advances in Statistics and Optimization, Editors: Khan, A. H. and Khan, R. Published by R.J Enterprises, New Delhi, India, 10-25.
5. Awad (2014). Review of Optimal Designs of Progressive Censoring Schemes submitted for publication.
6. Balakrishnan, N., Aggarwala, R., (2000). Progressive Censoring: Theory, Methods and Applications. Birkhäuser, Boston.
7. Balakrishnan, N. and Han, D., (2007). Optimal progressive Type-II censoring schemes for nonparametric confidence intervals of quantiles, Commun. Statist. Compu. Simul., 36, 1247-1262.

8. Balakrishnan, N., Burkschat, M., Cramer, E. and Hofmann, G., (2008). Fisher information based progressive censoring plans, *Comput. Statist. Data Anal.*, 53, 366–380.
9. Balakrishnan, N. and Cramer, E., (2014). *The Art of Progressive Censoring: Applications to Reliability and Quality*, Springer Science+Business Media, N.Y.
10. Burkschat, M., Cramer, E. and Kamps, U., (2006). “On Optimal Schemes in Progressive Censoring”, *Statist. Probab. Lett.*, 76, pp. 1032-1036.
11. Burkschat, M., Cramer, E. and Kamps, U., (2007). “Optimality Criteria and Optimal Schemes in Progressive Censoring”, *Comm. Statist. Theory Methods*, 36, pp. 1419-1431.
12. Burkschat, M., (2008a). On optimality of extremal schemes in progressive Type-II censoring, *J. Statist. Plan. Inference*, 138, 1647–1659.
13. Burkschat, M., (2008b). On optimal linear equivariant estimation under progressive censoring. *Statistics*, 42, 5, 383-392.
14. Cramer, E., (2002). A note on moments of progressively Type II censored order statistics, *Comm. Statist. Theory Methods*, 31, 1301-1307.
15. Cramer, E. and Bagh, C. (2011). Minimum and maximum information censoring plans in progressive censoring. *Communications in Statistics-Theory and Methods*, 40, 14, 2011, 2511-2527.
16. Gini, C., (1912). Variabilita' e Mutabilita', *Studi Economicoaguaridici della Facotta di Ginrisprudenza dell, Universite di Cagliari III, Parte II*.
17. Gupta, R. D. and Kundu, D., (2006). Comparison of Fisher information matrices of the Weibull and generalized exponential distributions, *Journal of Statistical Planning and Inference*, 136, 3130 - 3144.
18. Haj Ahmed, H. A. and Awad, A. M. (2009a). Optimal Two-stage Progressive Type-II Censoring with Exponential Life-time, *Dirasat, Pure Sciences*, 36,2, 134-144
19. Haj Ahmed, H. A. and Awad, A. M. (2009b). Optimality Criterion for Progressive Type-II Right Censoring Based on Awad Sup-entropy Measures , *Journal of Statistics*, 16, 12-27.
20. Havrda, J. and Charvat, F., (1967). “Quantification Method of Classification Processes: Concept of Structural a-Entropy”, *Kybernetika*, 3, 30-35.
21. Hofmann G., Cramer E., Balakrishnan N., and Kunert G., (2005) . An asymptotic approach to progressive censoring. *J. Stat. Plann. Inference*, 130, 207–227.
22. Kamps U. (1995). A concept of generalized order statistics. *J Stat Plann Inference* 48:1–23.
23. Ng, H. K. T., Chan, C. S. and Balakrishnan, N., (2002). Estimation of parameters from progressively censored data using EM algorithm, *Computational Statistics and Data Analysis*, 39, 371 – 386.
24. Ng HKT, Chan PS, Balakrishnan N (2004). Optimal progressive censoring plans for the Weibull distribution. *Technometrics* 46:470–481.

25. Oja, H., (1981). On location, scale, skewness and kurtosis of univariate distributions. *Scand. J. Statist.* 8, 154-168.
26. Pradhan, B. and Kundu, D., (2009). On progressively censored generalized exponential distribution, *Test*, 18, 497 - 515.
27. Renyi, A., (1961), "On Measures of Entropy and Information", *Proc. 4th Berk. Symp. Math. Statist. and Probl.*, University of California Press, Vol. 1, 547-561.
28. Shannon, C.E., (1948), "A Mathematical Theory of Communication", *Bell Syst. Tech. J.*, 27, 379-423, pp. 623-656.
29. Sharma, B.D. and Mital, D.P. (1975), "New Nonadditive Measures of Inaccuracy", *J. Math Sci.*, 10, 122-133.
30. Soliman, A. A. (2005). Estimation of parameters of life from progressively censored data using Burr-XII model. *IEEE Transactions on Reliability*, 54,1, 34-42.
31. Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics, *J. Statist. Phys.*, vol. 52, pp. 479-487.

Appendix

Table 1: Shannon entropy in vicinity of $\theta^* = 0.508329$; $\varepsilon = 0.1, m = 3, n = 8$

R	$H(R; \theta^* - \varepsilon)$	$H(R; \theta^*)$	$H(R; \theta^* + \varepsilon)$
{0, 0, 5}	-2.15633	-2.41488	-2.68608
{0, 1, 4}	-2.05564	-2.29813	-2.55855
{0, 2, 3}	-1.95494	-2.17335	-2.4176
{0, 3, 2}	-1.87135	-2.0496	-2.26691
{0, 4, 1}	-1.87405	-1.97201	-2.13542
{0, 5, 0}	-2.4054	-2.26248	-2.26419
{1, 0, 4}	-2.01811	-2.23766	-2.48268
{1, 1, 3}	-1.91741	-2.11288	-2.34173
{1, 2, 2}	-1.83381	-1.98913	-2.19104
{1, 3, 1}	-1.83652	-1.91154	-2.05955
{1, 4, 0}	-2.36787	-2.20201	-2.18832
{2, 0, 3}	-1.89836	-2.06171	-2.269
{2, 1, 2}	-1.81476	-1.93796	-2.1183
{2, 2, 1}	-1.81746	-1.86037	-1.98681
{2, 3, 0}	-2.34882	-2.15083	-2.11559
{3, 0, 2}	-1.83652	-1.91154	-2.05955
{3, 1, 1}	-1.83922	-1.83394	-1.92805
{3, 2, 0}	-2.37058	-2.12441	-2.05683
{4, 0, 1}	-1.95971	-1.87413	-1.91435
{4, 1, 0}	-2.49106	-2.1646	-2.04312
{5, 0, 0}	-2.90193	-2.41488	-2.18561
R Maximum	{2,1,2}	{3,1,1}	{4,0,1}
H maximum	-1.81476	-1.83394	-1.91435
R minimum	{5,0,0}	{5,0,0} & {0,0,5}	{0,0,5}
H minimum	-2.90193	-2.41488	-2.68608

Table 2: Pairs of AH-equivalent Schemes

n	m	R	S	n	m	R	S
7	3	(2,0,2)	(0,3,1)	10	4	(3,0,1,2)	(0,3,2,1)
8	3	(3,0,2)	(1,3,1)			(3,1,0,2)	(0,4,1,1)
	4	(0,2,0,2)	(0,0,3,1)			(3,1,1,1)	(0,5,0,1)
9	3	(2,1,3)	(0,4,2)			(3,1,2,0)	(0,5,1,0)
		(4,0,2)	(2,3,1)			(1,3,0,2)	(1,1,3,1)
	4	(0,3,0,2)	(0,1,3,1)			(2,2,0,2)	(2,0,3,1)
		(1,2,0,2)	(1,0,3,1)		5	(0,0,3,0,2)	(0,0,1,3,1)
	5	(0,0,2,0,2)	(0,0,0,3,1)			(0,1,2,0,2)	(0,1,0,3,1)
		(2,0,0,0,2)	(0,2,0,1,1)			(1,0,2,0,2)	(1,0,0,3,1)
		(2,0,0,1,1)	(0,2,1,0,1)			(3,0,0,0,2)	(1,2,0,1,1)
		(2,0,0,2,0)	(0,2,1,1,0)			(3,0,0,1,1)	(1,2,1,0,1)
						(3,0,0,2,0)	(1,2,1,1,0)
10	3	(3,1,3)	(1,4,2)		6	(0,0,0,2,0,2)	(0,0,0,0,3,1)
		(5,0,2)	(3,3,1)			(0,2,0,0,0,2)	(0,0,2,0,1,1)
	4	(0,2,1,3)	(0,0,4,2)			(0,2,0,0,1,1)	(0,0,2,1,0,1)
		(0,4,0,2)	(0,2,3,1)			(0,2,0,0,2,0)	(0,0,2,1,1,0)
		(3,2,0,1)	(0,2,4,0)				

Table 3: Values of θ at which " $\Delta HRC = H(R) - H(C)$ " is minimum and the efficiency of the scheme is greater than $m/n = 0.50$

Schemes	ΔHRC	θ	$H(C)$	$H(R)$	Efficiency
{0,0,0,0,5}	7.27812	1.64563	-16.1624	-8.88423	0.549687
{0,0,0,1,4}	7.44015	1.63897	-16.1465	-8.70633	0.539209
{0,0,1,0,4}	7.56516	1.62944	-16.1239	-8.55871	0.53081
{0,0,0,2,3}	7.63269	1.62897	-16.1227	-8.49004	0.526588
{0,1,0,0,4}	7.66571	1.61873	-16.0985	-8.43281	0.523825
{1,0,0,0,4}	7.74905	1.60762	-16.0723	-8.32328	0.517864
{0,0,1,1,3}	7.75752	1.61944	-16.1002	-8.34268	0.518172
{0,1,0,1,3}	7.85786	1.60873	-16.0749	-8.21708	0.511173
{0,0,0,3,2}	7.86895	1.6123	-16.0834	-8.2144	0.510739
{0,0,2,0,3}	7.89851	1.60611	-16.0688	-8.17028	0.508457
{1,0,0,1,3}	7.94099	1.59762	-16.0489	-8.10788	0.505199
{0,0,1,2,2}	7.99348	1.60278	-16.061	-8.06748	0.502304
{0,1,1,0,3}	7.99857	1.5954	-16.0437	-8.0451	0.50145

