

Testing exponentiality based on Progressively Type-I interval censored data

Huda H. Qubbaj*

Department of Mathematics, The University of Jordan, Jordan

Abstract In this paper we propose non-parametric estimates for the information measure entropy when a progressively Type-I interval censored data is available. Different non-parametric approaches are used for deriving the estimates. Entropy-based tests of exponentiality are proposed. The critical values and the power values of the proposed tests are simulated and studied under various alternatives. Real life data sets are presented and analysed.

Keywords Entropy; Non-parametric statistics; Type-I Interval Censoring; Testing exponentiality; Monte Carlo simulation.

DOI: 10.19139/soic-2310-5070-2394

1. Introduction

Shannon entropy [1] of a random variable (r.v.) X whose probability density function (pdf) $f(x)$ and cumulative distribution function (cdf) $F(x)$, is defined as:

$$H(X) = - \int_{R_X} f(x) \log(f(x)) dx, \quad (1)$$

where R_X denotes the support of the r.v. X .

For more details on entropy the reader can see [4,13,20,21,25]. Also [11,12] introduced nonparametric estimates for entropy based on progressively Type-II censored data. Furthermore, the estimation problem of certain entropy measures for particular distributions under a specific type of censoring have been discussed in literature. For example, [2] studied Entropy Estimation of Inverse Weibull Distribution under Improved Adaptive Progressively Type-II Censoring. In literature there are different approaches and versions of entropy estimations that provide a diverse toolkit for researchers to choose from depending on the specific application and the data characteristics.

However, estimation of the entropy measure under progressively Type-I interval censored data have not been considered so far in the Literature. Accordingly, our main objective in this paper is to use the developed different methods for estimating the entropy measures under the progressive Type-I censoring set-up in testing exponentiality.

[5] introduced progressive Type-I censoring as an extension of Type-I censoring, where in a progressively Type-I censored life test on n items, progressive censoring is carried out at the prefixed censoring times $t_1 < t_2 < \dots < t_k$. That is, at the i^{th} censoring time t_i , R_i items are randomly removed from the experiment, $1 \leq i \leq k-1$, with the restriction $R_1 + \dots + R_{k-1} \leq n-l$, $l \in \{0, \dots, n\}$. Then at the k^{th} censoring time t_k , all remaining items are removed from the life test if there are any left. In many practical situations lifetimes of units placed on a test are observed within Intervals, where this censoring scheme is called Interval censoring.

*Correspondence to: Huda H. Qubbaj (Email: hudaqubbaj.ctc@gmail.com). Department of Mathematics, University of Jordan. Queen Rania street, Amman, Jordan.

[1] initially discussed progressive type-I interval censoring in literature and studied an exponential distribution using this censoring.

Since then this censoring scheme has attracted attention among researchers. Progressive type-I interval censoring can be briefly described as follows: Suppose n identical items are placed simultaneously on life testing at time $t_0 = 0$, where inspection is at m pre-fixed censoring times $t_1 < t_2 < \dots < t_m$, and where t_m is the scheduled time to terminate the experiment and m is pre-fixed number of stops. For $i = 1, 2, \dots, m$, let k_i be the number of failures in the interval $(t_{i-1}, t_i]$. Let S_i be the number of the surviving items at t_i and R_i be the number of removed items at t_i . In this censoring scheme, k_i and S_i are random numbers while R_i is the number of remaining items, which is also a random number. At the 1st inspection time t_1 , we observe k_1 failures, then R_1 surviving items are randomly withdrawn from the remaining items $n - k_1$. One can see that after this step, the number of remaining items is $(n - k_1 - R_1)$. Now, after time t_1 and at the 2nd inspection time t_2 , we observe k_2 failures where R_2 items are randomly removed from $(n - k_1 - k_2 - R_1)$ items. Lastly, at the m^{th} inspection time (the last inspection time), we observe k_m failures and all remaining $(n - \sum_{i=1}^m k_i - \sum_{i=1}^{m-1} R_i)$ items are immediately removed from the experiment.

The observed progressive Type-I interval censored data can be represented as:

$$\{(k_i, R_i, t_i), i = 1, 2, \dots, m\}.$$

The associated likelihood function under the progressive type-I interval censoring is given by:

$$L(\theta) \propto \prod_{i=1}^m [F(t_i; \theta) - F(t_{i-1}; \theta)]^{k_i} [1 - F(t_i; \theta)]^{R_i}. \quad (2)$$

Note that R_i should not be greater than S_i , where the values of R_i for $i = 1, 2, \dots, m$ are determined based on pre-specified removal proportions $q_1, q_2, \dots, q_{k-1}$ and $q_m = 1$, such that $R_i = [S_i q_i]$, for $i = 1, 2, \dots, k - 1$, where symbol $[b]$ is the greatest integer less than or equal to b . Progressive type-I interval censoring approach has been considered by different authors in the literature including [18], Lio et. al.[16], [24], Du et. al.[8], Al otaibi et. al.[3] and Qubbaj et. al.[14,15].

The rest of the paper is organized as follows: Non-parametric estimation for entropy measure based on progressive Type-I interval censoring are developed in Section 2. In section 3, we presented critical values of the proposed tests and then powers of the tests against different alternatives are computed by Monte Carlo simulation. In section 4, we illustrated the proposed tests by real-life data example. Conclusions were presented in section 5. Finally we ended section 6 with limitations and future directions.

The next theorem considers Type-I interval censoring using an underlying lifetime distribution, namely the uniform.

Theorem 1.1

Let $U_{i:m:n} = F(t_i)$, $i = 1, 2, \dots, m$ denote a progressively Type-I interval censoring sample obtained from the uniform $(0, 1)$ distribution, assuming the sample size is n with progressive Type-I interval censored data $\{(k_i, R_i, t_i), i = 1, 2, \dots, m\}$. Let

$$U_{i:m:n} = 1 - \prod_{j=m-i+1}^m V_j,$$

where,

$$V_1 = \frac{1 - U_{m:m:n}}{1 - U_{m-1:m:n}}, V_2 = \frac{1 - U_{m-1:m:n}}{1 - U_{m-2:m:n}}, \dots, V_m = 1 - U_{1:m:n}, \quad (3)$$

are all independent identically distributed (iid) r.v.'s. Then

$$V_i \stackrel{d}{=} \text{Beta} \left(i + \sum_{j=m-i+2}^m k_j + \sum_{j=m-i+1}^m R_j, k_{m-i+1} + 1 \right), \quad i = 1, 2, \dots, m. \quad (4)$$

The proof was detailed in [14] so it is omitted.

Corollary 1.2

As a result of Theorem 1.1, we find

$$E(U_{i:m:n}) = 1 - \prod_{j=m-i+1}^m \gamma_j, \quad (5)$$

where,

$$\gamma_i = \frac{i + \sum_{j=m-i+2}^m k_j + \sum_{j=m-i+1}^m R_j}{1 + i + \sum_{j=m-i+1}^m k_j + R_j},$$

such that $\gamma_j = \gamma_1$ if $j \leq 1$ and $\gamma_j = \gamma_m$ if $j \geq m$ provided that $\sum_{j=m+1}^m k_j = 0$.

2. Non-parametric Entropy Estimates

This section develops non-parametric estimates for the entropy measures based on progressively Type-I interval censored samples. It is of importance here to mention that for a random variable (r.v.) T , entropy measure $H(T)$ is expressed as:

$$H(T) = \int_0^1 \log \left(\frac{d}{dp} F^{-1}(p) \right) dp \quad (6)$$

2.1. Moments Approximation Method

The first entropy estimate will be obtained by using the difference operator that was proposed by [26] for estimating the entropy. This method is based on the following fact:

$$\frac{d}{dp} F^{-1}(p) \approx \frac{T_{i+w:m:n} - T_{i-w:m:n}}{F(T_{i+w:m:n}) - F(T_{i-w:m:n})}, \quad (7)$$

where the window size $w \leq m/2$, also $T_{i:m:n} = T_0 = 0$ if $i < 1$, while $T_{i:m:n} = T_{m:m:n}$ if $i > m$. Then $T_0 = 0 \leq T_{1:m:n} \leq T_{2:m:n} \leq \dots \leq T_{m:m:n}$ are progressively Type-I interval censoring times of size m , which are pre-fixed. In the interval $[T_{i-1}, T_i]$, we observe a random number of failures, say k_i , then R_i surviving units are immediately removed from the remaining $(n - \sum_{j=1}^i k_j - \sum_{j=1}^{i-1} R_j)$ items.

Following the lines of [26] and by using (2.11), entropy $H(T)$ can be approximated as:

$$H(T) = \frac{1}{m} \sum_{i=1}^m \log \left(\frac{T_{i+w:m:n} - T_{i-w:m:n}}{E(U_{i+w:m:n}) - E(U_{i-w:m:n})} \right). \quad (8)$$

First estimate of the entropy is

$$\hat{H}_1 = \frac{1}{m} \sum_{i=1}^m \log \left(\frac{T_{i+w:m:n} - T_{i-w:m:n}}{\prod_{j=m-(i-w)+1}^m \gamma_j - \prod_{j=m-(i+w)+1}^m \gamma_j} \right). \quad (9)$$

Proposition 2.1

Let $Y = aT + b$, $a > 0$. Then $\hat{H}_1^Y = \log(a) + \hat{H}_1^T$.

Proof

The result follows since for all i , $\gamma_i^Y = \gamma_i^T$. □

Proposition 2.2

Estimate $\hat{H}_1$ is consistent estimats for H i.e.

$$\hat{H}_1 \xrightarrow{p} H.$$

as $n \rightarrow \infty$, $m \rightarrow \infty$, and $m/n \rightarrow 0$.

Proof

Since when $m \rightarrow n$, the progressively Type-I interval censored sample becomes the complete sample. Then the estimate, $\hat{H}_1$ converge to the estimates proposed by [26], which is a consistent estimate of H . $\square$

2.2. Linear Approximation Method

The second estimate for entropy $H(T)$ is proposed following the steps of [6] by noticing that the quantity:

$$\frac{F(T_{i+w:m:n}) - F(T_{i-w:m:n})}{T_{i+w:m:n} - T_{i-w:m:n}}, \quad (10)$$

represents the slope of a straight line joining the following two points

$$(T_{i-w:m:n}, F(T_{i-w:m:n})) \text{ and } (T_{i+w:m:n}, F(T_{i+w:m:n})).$$

This estimation approach is based on estimating the function $F(T_j)$ by a local linear model on $(T_{i-w:m:n}, T_{i+w:m:n})$ by using $2w + 1$ ordered pairs:

$$F(T_j) = U_j = a + bT_j + \epsilon, j = i - w, \dots, i + w. \quad (11)$$

On the other hand, slope in (11) can be approximated by b in (12), which can be estimated by the least squares method using $(2w + 1)$ ordered pairs as follows:

$$b = \frac{S_{TU}}{S_T^2} = \frac{\sum_{j=i-w}^{i+w} (T_{j:m:n} - \bar{T}_{(i)})(U_{j:m:n} - \bar{U}_{(i)})}{\sum_{j=i-w}^{i+w} (T_{j:m:n} - \bar{T}_{(i)})^2}, \quad (12)$$

where,

$$\bar{T}_{(i)} = \frac{1}{2w+1} \sum_{j=i-w}^{i+w} T_{j:m:n}, \text{ and } \bar{U}_{(i)} = \frac{1}{2w+1} \sum_{j=i-w}^{i+w} U_{j:m:n}.$$

Now, by replacing $U_{j:m:n}$ by its expected value, we get

$$\hat{U}_{(i)} = \frac{1}{2w+1} \sum_{j=i-w}^{i+w} \left(1 - \prod_{k=m-j+1}^m \gamma_k \right).$$

Similarly, we consider the slope of the linear regression of T on F as follows:

$$b_h = \frac{S_{TU}}{S_U^2} = \frac{\sum_{j=i-w}^{i+w} (T_{j:m:n} - \bar{T}_{(i)})(U_{j:m:n} - \bar{U}_{(i)})}{\sum_{j=i-w}^{i+w} (U_{j:m:n} - \bar{U}_{(i)})^2}, \quad (13)$$

Using Eq.(13), and replacing $(U_{j:m:n} - \bar{U}_{(i)})$ by $(\hat{U}_{j:m:n} - \hat{\bar{U}}_{(i)})$ in it, a second estimate for $H(T)$ is introduced as follows:

$$\hat{H}_2 = \frac{1}{m} \sum_{i=1}^m \log \left[\frac{\sum_{j=i-w}^{i+w} (T_{j:m:n} - \bar{T}_{(i)}) \left(\frac{\sum_{j=i-w}^{i+w} \prod_{k=m-j+1}^m \gamma_k}{2w+1} - \prod_{k=m-j+1}^m \gamma_k \right)}{\sum_{j=i-w}^{i+w} \left(\frac{\sum_{j=i-w}^{i+w} \prod_{k=m-j+1}^m \gamma_k}{2w+1} - \prod_{k=m-j+1}^m \gamma_k \right)^2} \right]. \quad (14)$$

Proposition 2.3

Let $Y = aT + b$, $a > 0$. Then $\hat{H}_2^Y = \log(a) + \hat{H}_2^T$.

Proof

The result follows since for all i , $\gamma_i^Y = \gamma_i^T$. □

Proposition 2.4

Estimate $\hat{H}_2$ is consistent estimates for H , i.e.

$$\hat{H}_2 \xrightarrow{p} H.$$

as $n \rightarrow \infty$, $m \rightarrow \infty$ and $m/n \rightarrow 0$.

Proof

Proof is obvious by [6] and so it is omitted. □

2.3. Kernel-Based Method

Here, the entropy $H(T)$ can also be represented by the form $-E_f(\log f(t))$. Accordingly an estimate of H is

$$\hat{H}_3 = -\frac{1}{m} \sum_{i=1}^m \log \left(\hat{f}(t_{i:m:n}) \right). \quad (15)$$

where $\hat{f}(t_{i:m:n})$ is estimated by the kernel function K as

$$\hat{f}(t_{i:m:n}) = \frac{1}{mh} \sum_{j=1}^m K \left(\frac{t_{i:m:n} - t_{j:m:n}}{h} \right), \quad (16)$$

and d is the bandwidth such that $d > 0$, also d is called the smoothing parameter or window width by some authors, K is the kernel function that is non-negative, smooth and symmetric function which satisfies the conditions [7]:

$$\int K(x)dx = 1, \quad \text{and} \quad \int xK(x)dx = 0.$$

This estimate is proposed assuming the Kernel function is the standard normal density function; due its convenient mathematical properties the normal Kernel is frequently used

$$K(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$

There are several choices for the bandwidth d . The bandwidth d is chosen here to equal $1.06Sm^{-\frac{1}{5}}$ which is the optimal choice for d i.e. the bandwidth that minimizes the mean integrated square error [23], where S is the sample standard deviation and m is the number of points.

Proposition 2.1

Let $Y = aT + b$, $a > 0$. Then $\hat{H}_3^Y = \log(a) + \hat{H}_3^T$.

Proof

The results can be obtained upon using the transformation of the r.v t ; let $t = \frac{y-b}{a}$, hence $|\frac{dt}{dy}| = \frac{1}{|a|}$, thus, $f_y(y) = \frac{1}{|a|} f_t(\frac{y-b}{a}) = \frac{1}{|a|} f_t(t)$. □

Proposition 2.2

$\hat{H}_3$ is a consistent estimate for H , i.e.

$$\hat{H}_3 \xrightarrow{p} H.$$

as $m \rightarrow n$, $n \rightarrow \infty$, $w \rightarrow \infty$ and $w/m \rightarrow 0$.

Proof

We have,

$$\hat{H}_3 = -E(\log(\hat{f}(x))) \xrightarrow{p} -E(\log(f(x))) = H_3.$$

Hence, the result. $\square$

Several properties of the suggested estimates were introduced in [14,15].

3. Simulation Study

3.1. Introduction to the tests

In this section, we investigate entropy-based tests of exponentiality of a progressively Type-I interval censored sample, based on the proposed estimates of entropy given in section 2. Suppose $X_1, \dots, X_n$ is a random sample from a continuous probability distribution function F with density f over a non-negative support with mean $\mu < \infty$.

We are interested in testing the null hypothesis

$$H_0 : f(x) = f_0(x) = \lambda \exp(-\lambda x), \text{ for all } x \in (0, \infty)$$

against the alternative hypothesis

$$H_1 : f(x) \neq f_0(x), \text{ for some } x \in (0, \infty). \text{ Where } \lambda = \frac{1}{\mu} \text{ is unspecified.}$$

Following [6,9,10,19], we propose entropy-based test statistics based on complete samples as:

$$TH_i = -\hat{H}_i, \quad i = 1, 2, 3. \quad (17)$$

Also, the exponentiality is rejected for large values of the test statistic TH_U^i .

Table 1. Progressive interval censoring schemes used in Monte Carlo simulation study

Scheme No.	$(t_1, \dots, t_4)$	$(q_1, \dots, q_4)$
1	0.1, 0.5, 0.7, 0.9	0, 0, 0.25, 1
2	0.1, 0.5, 0.7, 0.9	0, 0.25, 0.25, 1
3	0.1, 0.5, 0.7, 0.9	0.25, 0, 0, 1
4	0.1, 0.5, 0.7, 0.9	0.2, 0.2, 0.2, 1
5	0.1, 0.5, 0.7, 0.9	0, 0, 0, 1

3.2. Critical Values

The proposed GOF tests were studied via Monte Carlo simulation based on censoring schemes given in Table (1). Simulated critical values are obtained under several progressive Type-I interval censoring schemes assuming different sample sizes $n = 20, 50$ and 100 for $m = 4$ at (0.05) level of significance with (10000) replications, these critical values are presented in Table (2). The computations were made using Mathematica packages.

Table 2. Critical values of the GOF statistics $T\hat{H}_i$ ($i = 1, 2, 3$) at significance level $\alpha = 0.05$ for $m = 4$

n	Scheme	$T\hat{H}_1$	$T\hat{H}_2$	$T\hat{H}_3$
20	1	0.109	0.141	0.897
	2	0.121	0.143	0.930
	3	0.076	0.105	0.797
	4	0.039	0.078	0.994
	5	0.109	0.136	0.750
50	1	0.229	0.261	0.562
	2	0.229	0.259	0.570
	3	0.171	0.212	0.474
	4	0.173	0.213	0.622
	5	0.184	0.221	0.493
100	1	0.298	0.332	0.522
	2	0.357	0.391	0.597
	3	0.237	0.278	0.457
	4	0.276	0.311	0.579
	5	0.289	0.325	0.413

Table 3. Monte Carlo power estimates at significance level $\alpha = 0.05$ for $m = 4$: Monotonic decreasing hazard alternatives

Alternative	Scheme	n	$T\hat{H}_1$	$T\hat{H}_2$	$T\hat{H}_3$
Gamma(0.5,1)	1	20	0.257	0.237	0.050
	2		0.277	0.257	0.089
	3		0.168	0.158	0.109
	4		0.228	0.228	0.099
	5		0.267	0.257	0.079
$\chi^2(1)$	1	20	0.079	0.069	0.059
	2		0.069	0.050	0.089
	3		0.050	0.040	0.109
	4		0.050	0.050	0.089
	5		0.069	0.050	0.119
Gamma(0.5,1)	1	50	0.297	0.307	0.287
	2		0.208	0.208	0.564
	3		0.168	0.178	0.614
	4		0.178	0.208	0.287
	5		0.208	0.208	0.584
$\chi^2(1)$	1	50	0.069	0.069	0.158
	2		0.089	0.089	0.188
	3		0.020	0.020	0.356
	4		0.020	0.020	0.208
	5		0.000	0.000	0.238
Gamma(0.5,1)	1	100	0.248	0.248	0.970
	2		0.396	0.406	0.901
	3		0.178	0.178	0.990
	4		0.248	0.267	0.901
	5		0.337	0.317	0.980
$\chi^2(1)$	1	100	0.039	0.040	0.673
	2		0.049	0.045	0.584
	3		0.039	0.040	0.812
	4		0.049	0.045	0.455
	5		0.039	0.030	0.792

3.3. Power Comparisons

The statistical powers of the proposed tests are studied via simulation using Monte Carlo method under various censoring schemes (shown in Table 1), assuming a set of alternative probability distributions at 0.05

level of significance. Under each alternative, we have generated 10000 samples of moderate and large sizes $n = 20, 50$ and 100 for $m = 4$ i.e four stopping times. In this study we considered the following alternatives according to the type of the hazard function:

- Group I alternative (Monotonic decreasing hazard functions): Gamma (shape parameter: 0.5) and Chi-square (degree of freedom,1).
- Group II alternative (Monotonic increasing hazard functions):Beta (shape parameter:1) and Uniform.
- Group III alternative (Non-monotonic hazard functions): Log normal (shape parameters 0.5) and Beta (shape parameter:0.5).

The statistical powers are displayed in Tables (3-5). The test statistic $T\hat{H}_3$ outperformed the other test statistics for the monotonic decreasing hazard function alternatives. However, the test statistic $T\hat{H}_2$ had the best power than other test statistics for the monotonic increasing and non-monotonic hazard function alternatives.

Table 4. Monte Carlo power estimates at significance level $\alpha = 0.05$ for $m = 4$: Monotonic increasing hazard alternatives

Alternative	Scheme	n	$T\hat{H}_1$	$T\hat{H}_2$	$T\hat{H}_3$
Beta(1,2)	1	20	0.743	0.792	0.069
	2		0.703	0.723	0.089
	3		0.634	0.633	0.089
	4		0.713	0.723	0.040
	5		0.753	0.752	0.119
Uniform(0,1)	1	20	0.743	0.762	0.000
	2		0.792	0.792	0.019
	3		0.832	0.822	0.000
	4		0.643	0.644	0.000
	5		0.881	0.881	0.020
Beta(1,2)	1	50	0.970	0.970	0.218
	2		0.921	0.941	0.455
	3		0.931	0.960	0.307
	4		0.940	0.941	0.178
	5		0.931	0.960	0.248
Uniform(0,1)	1	50	0.980	0.990	0.000
	2		0.960	0.990	0.000
	3		0.951	0.960	0.000
	4		0.960	0.970	0.000
	5		1.000	1.000	0.010
Beta(1,2)	1	100	1.000	1.000	0.396
	2		1.000	1.000	0.554
	3		0.960	0.980	0.554
	4		0.990	1.000	0.584
	5		0.970	0.980	0.861
Uniform(0,1)	1	100	1.000	1.000	0.010
	2		1.000	1.000	0.010
	3		1.000	1.000	0.000
	4		0.990	1.000	0.000
	5		1.000	1.000	0.050

4. Real Data Examples

Example 1: The following data represents failure times (in minutes) for an insulating fluid between two electrodes subjected to a voltage of 34 kV. [17]:

0.19, 0.78, 0.96, 1.31, 2.78, 3.16, 4.15, 4.67, 4.85, 6.50,
7.35, 8.01, 8.27, 12.06, 31.75, 32.52, 33.91, 36.71, 72.89.

Let us consider this progressively Type-I interval censored sample of size $n = 19$, for the suggested censoring scheme we have $m = 4$ and $w = 2$, the stopping times are chosen to be (0,1,5,15,25) and the

Table 5. Monte Carlo power estimates at significance level $\alpha = 0.05$ for $m = 4$: Non-monotonic hazard alternatives

Alternative	Scheme	n	$T\hat{H}_1$	$T\hat{H}_2$	$T\hat{H}_3$
Log Normal(0,0.5)	1	20	0.060	0.120	0.059
	2		0.139	0.139	0.049
	3		0.020	0.020	0.089
	4		0.010	0.020	0.030
	5		0.030	0.059	0.139
Beta(0.5,1)	1	20	0.772	0.782	0.010
	2		0.772	0.772	0.000
	3		0.644	0.624	0.020
	4		0.654	0.673	0.030
	5		0.772	0.762	0.029
Log Normal(0,0.5)	1	50	0.079	0.119	0.059
	2		0.119	0.118	0.069
	3		0.118	0.149	0.208
	4		0.079	0.079	0.059
	5		0.059	0.138	0.128
Beta(0.5,1)	1	50	0.920	0.931	0.039
	2		0.891	0.901	0.149
	3		0.881	0.901	0.257
	4		0.851	0.861	0.059
	5		0.921	0.931	0.158
Log Normal(0,0.5)	1	100	0.168	0.188	0.099
	2		0.218	0.228	0.010
	3		0.208	0.238	0.228
	4		0.198	0.208	0.040
	5		0.198	0.277	0.426
Beta(0.5,1)	1	100	0.980	0.990	0.594
	2		0.960	0.990	0.505
	3		0.980	0.980	0.683
	4		0.980	0.980	0.485
	5		0.990	0.990	0.455

applied proportions are $(0.25, 0, 0.25, 1)$, it is worth mentioning that this scheme was chosen according to the domain-specific reasoning. The values of the entropy test statistics and the corresponding critical values of the tests are computed and presented in Table (6), which proves that the observed progressively Type-I interval censored sample is from an exponential distribution.

Table 6. Values of the test statistics and the critical values in Example 1

$T\hat{H}_i$	Value of $T\hat{H}_i$	Critical value
$T\hat{H}_1$	3.509	0.1811
$T\hat{H}_2$	3.507	0.1788
$T\hat{H}_3$	2.973	0.1673

Example 2:

The following data are obtained based on inter-occurrence times in days for fatal accidents suffered by scheduled large planes in the USA for years from 1983 to 1998. (data from NTSB):

2, 5, 7, 10, 11, 13, 14, 16, 17, 22, 22, 22, 22, 35, 36, 41, 50, 53, 53, 56, 60, 61, 63, 63, 65, 68, 70, 91, 98, 112, 116, 117, 125, 125, 127, 128, 143, 143, 148, 150, 151, 158, 162, 194, 216, 223, 236, 244, 253, 310, 426, 454.

Here, we will consider a progressively Type-I interval censored sample of size $n = 52$, for the suggested censoring scheme we have $m = 4$ and $w = 2$, the stopping times are chosen to be $(0, 20, 60, 150, 500)$ and the applied proportions are $(0, 0, 0.25, 1)$, moreover this scheme was chosen according to the domain-specific

reasoning. The values of the entropy test statistics and the corresponding critical values of the tests are computed and presented in Table (7), and we conclude that the inter-occurrence times of fatal accidents suffered by scheduled large planes in the USA in period (1983 - 1998) are exponentially distributed.

Table 7. Values of the test statistics and the critical values in Example 2

$T\hat{H}_i$	Value of $T\hat{H}_i$	Critical value
$T\hat{H}_1$	-6.277	0.925
$T\hat{H}_2$	-6.327	0.933
$T\hat{H}_3$	5.487	0.874

5. Conclusions

In this paper, Non-parametric based methods involving moments, linear and Kernel-based have been discussed. Entropy-based tests of exponentiality under progressively Type-I interval censored data were proposed using these estimates. Monte Carlo simulation showed that the proposed entropy-based tests under progressively type-I interval censoring perform well and provide satisfactory powers for testing exponentiality with different alternative hazard functions. By real data examples, we have shown that the proposed tests are applicable.

Limitations and Future Directions

The findings of this study have to be seen in light of some limitations such as: issues with the selection of the schemes, effect of sample size, Lack of previous research studies on GOF based on type-I interval censored data, the exclusive and only choice of the window size w and Kernel Bandwidth Sensitivity.

Although we have obtained results of testing exponentiality for progressively Type-I interval censored data, this research can be applied for other censoring techniques as: hybrid or generalized progressive censoring. Another research line to be considered is the entropy based GOF testing other distributions. Therefore, additional work on entropy estimation as well as entropy- based GOF tests are needed along this path.

Acknowledgement

We would like to thank the editor and the referees for their valuable comments.

REFERENCES

1. R. Aggarwala, *Progressive interval censoring: some mathematical results with applications to inference*, Commun Stat: Theory Methods, 30 (8–9):1921–1935, 2001.
2. F. Alam, and M. Nasser, *On Entropy Estimation of Inverse Weibull Distribution under Improved Adaptive Progressively Type-II Censoring with Applications*, Axioms, 12(8), 751, 2023.
3. R. Alotaibi, H. Rezk, S. Dey, and H. Okasha, *Bayesian estimation for Dagum distribution based on progressive type I interval censoring*, PLOS ONE, 16(6), 2021.
4. A. M. Awad, and A. Alawneh, *Application of entropy of a life time model*, IMA J. Math. Control Inf., 4,143–147, 1987.
5. AC. Cohen, *Progressively censored sample in life testing*. Technometrics, 5, 327–339, 1963.
6. J.C. Correa, *A new estimator of entropy*, Communications in Statistics–Theory and Methods, 24, 2439–2449, 1995.
7. Yu.G. Dmitriev, F.P. Tarasenko, *On estimation of functionals of the probability density function and its derivatives*. Teor. Veroyatnost. i Primenen., 18(3), 662–668, 1973.

8. Y. Du, Y. Guo and W. Gui, *Statistical Inference for the Information Entropy of the Log-Logistic Distribution under Progressive Type-I Interval Censoring Schemes*, *Symmetry*, 10, 445, 2018.
9. N. Ebrahimi, K. Pflughoeft, and E. S. Soofi, *Two measures of sample entropy*. *Statistics and Probability Letters*, 20(3), 225-234, 1994.
10. J. Jarrah, HA. Noyghabi, *Testing exponentiality using different entropy estimates based on type II censored data: a monte carlo power comparison*. *International Journal of Industrial Engineering*, 24(5), 556-571, 2017.
11. R. Hazebe, M. Raqab, and H. Bayoud, *Non-parametric estimation of the extropy and the entropy measures based on progressive type-II censored data with testing uniformity*. *Journal of Statistical Computation and Simulation*, 91 (11), 2178-2210, 2021a.
12. R. Hazebe, H. Bayoud, and M. Raqab, *Kernel and CDF-Based Estimation of Extropy and Entropy from Progressively Type-II Censoring with Application for Goodness of Fit Problems*. *Stochastic and Quality Control*, 36 (1), 73-83, 2021b.
13. O. A. Kittaneh, M. A. Khan, M. Akbar, and H. Bayoud, *Average entropy: a new uncertainty measure with application to image segmentation*. *The American Statistician*, 70 (1), 18-24, 2016.
14. H. Qubbaj, H. Bayoud, and H. Hilow, *Extropy and entropy estimation based on progressive Type-I interval censoring*, *Statistics in transition new series*, Vol. 25, No. 3, pp. 83-102, 2024.
15. H. Qubbaj, H. Bayoud, and H. Hilow, *Entropy and Entropy Estimation Based on Uniformly and Log-Normal Distributed Data*, *JJMS* 18, No. 1, 115-125, 2025.
16. YL. Lio, DG. Chen, and TR. Tasi, *Parameter estimations for generalized exponential distribution under progressive type-I interval censoring*. *Comput. Stat. Data Anal.*, 54, 1581-1591, 2011.
17. W. B. Nelson, *Life Data Analysis*, Wiley, 1982.
18. H. K. T. Ng, and Z. Wang, *Statistical estimation for the parameters of Weibulldistribution based on progressively Type-I interval censored sample*. *J. Stat. Comput.Simulat.*, 79(2):145-159, 2009.
19. HA. Noughabi, *A new estimator of entropy and its application in testing normality*. *Journal of Statistical Computation and Simulation*, 80 (10), 1151-1162, 2010.
20. M. Rao, Y. Chen, B. C. Vemuri, and F. Wang, *Cumulative residual entropy: a new measure of information*. *IEEE transactions on Information Theory*, 50 (6), 1220-1228, 2004.
21. A. Renyi, *On measures of entropy and information*, *proc. 4th Berkeley symp. Math. Stat. and Prob.*, 1, 547-561, 1961.
22. C. E. Shannon, *A mathematical theory of communications*, *Bell System Tech. J.*, 27(3): 379-423, 1948.
23. B.W. Silverman, *Density Estimation for Statistics and Data Analysis*. Chapman Hall, London, 1986.
24. S. Singh S, and YM. Tripathi, *Estimating the parameters of an inverse Weibull distribution under progressive type-I interval censoring*. *Stat. Pap.*, 59, 21-56, 2016.
25. C. Tsallis, *Possible generalization of Boltzmann-Gibbs statistics*. *J. Stat. Phys.*, 52, 470-487, 1988.
26. O. Vasicek, *A test for normality based on sample entropy*, *Journal of the Royal Statistical Society B*, 38, 54-59, 1976.