Journal of Statistical Modelling: Theory and Applications
Vol. 3, No. 2, 2022, pp. 1-14
Yazd University Press 2022

Research Paper

Optimality in Type I hybrid censoring with random sample size

Elham Basiri*, Elham Hossinzadeh
Department of Mathematics and Applications, Kosar University of Bojnord, Bojnord, Iran

Received: February 19, 2023/ Revised: May 13, 2023/ Accepted: May 16, 2023

Abstract: This paper considers the Type I hybrid censoring and investigates the optimal value for the sample size which is assumed as a truncated binomial random variable. Rayleigh distribution is considered for the lifetime distribution. Towards this end, various factors can be considered and the most important is the sampling cost criterion. Since the sample size is a random variable, the optimal parameter of the random sample size is determined so that the total cost of the test does not exceed a pre-determined value. Numerical calculations and a simulation study have been performed to evaluate the obtained results. Finally, the conclusion of the article is presented.

Keywords: Cost criterion; Optimal sample size; Type I hybrid censoring.
Mathematics Subject Classification (2010): 62F30, 62N01.

1 Introduction

The probability density function (pdf) and cumulative distribution function (cdf) of one-parameter Rayleigh distribution are given as

$$g_\sigma(y) = \frac{y}{\sigma^2}\exp\left(-\frac{y^2}{2\sigma^2}\right); \quad y>0, \sigma>0, \tag{1}$$

$$G_\sigma(y) = 1-\exp\left(-\frac{y^2}{2\sigma^2}\right); \quad y>0, \sigma>0, \tag{2}$$

respectively. The parameter σ is known as the scale parameter. Also, Rayleigh distribution is a special case of Weibull distribution. This model is useful in lifetime tests,

*Corresponding author: elhambasiri@kub.ac.ir

where the failure rate has linearity with time. Rayleigh distribution was discussed by some authors. See, for example, Asgharzadeh and Azizpour (2016), MirMostafaee et al. (2017) and Wu et al. (2023).

Many times in reliability issues, lifetime tests, survival analysis, and other applied fields, due to reasons such as limited time, lack of access to all units, or expensive units under the test, the experimenter cannot observe the exact life time of units under the test. Censored data occurs when complete information is not available on all units under test. Censoring can be performed in different ways depending on how the data is collected which the most famous are Type I and Type II censoring methods. If in a test with n units all units start working at time $t = 0$ and the experiment ends at a specified and pre-determined time T, then Type I censoring has occurred. Also, if in a test with n units the goal is to accurately record the time of the mth failure, $m \leq n$, the experiment continues until observing the mth failure and then stops. In such a case, the lifetime of the $(n - m)$ remaining units is censored. In this case, Type II censoring has happened.

A combination of Type I and II censoring schemes is known as hybrid censoring, which is divided into two types. Suppose n items are tested. For the first time, Epstein (1954) investigated a design in a survival experiment in which the experiment ends at the time $\tau = \min(Y_{m:n}, Y_0)$, where Y_0 and m are pre-determined values. Childs et al. (2003) called this censoring scheme as Type I hybrid censoring.

Optimal designing for different censoring methods is an issue usually done by using a suitable criterion function, and each of them may lead to a different optimal scheme. Many researchers have studied this issue. For example, the goal of research done by Ebrahimi (1988) was finding the optimal censoring scheme in Type I hybrid censoring which minimizes the expected cost function. Burkschat et al. (2006) and Burkschat et al. (2007) found optimal censoring schemes in the model of progressively Type II censoring by minimizing the variance and the covariance matrix of the best linear unbiased estimators for a location-scale family of distributions which includes exponential, uniform and Pareto distributions. Burkschat (2008) considered the best linear equivariant estimation in a particular location-scale family based on several progressively Type II censored samples and obtained the censoring schemes that minimize the mean squared error matrix of the estimators. Volterman et al. (2012) obtained the optimal progressive censoring scheme from the exponential distribution based on Pitman closeness criterion. It was shown that the optimal progressive censoring scheme is the usual Type II right censoring. Bhattacharya et al. (2014) proposed a cost function to determine the optimum life testing plans under hybrid censoring scheme. The optimum scheme was obtained by minimizing the total cost associated with the experiment. Basiri (2017) found the Bayesian point predictor for future order statistics from the one-parameter Rayleigh distribution. Then, by considering two criteria as mean squared prediction error and total cost of the test, the optimal number of failures was determined. Mishra (2020) obtained the optimal one-step censoring schemes under the entropy criterion. The optimal Bayesian sampling plan for the two parameter exponential distribution under Type I hybrid censoring scheme was studied by Prajapat et al. (2021). Sen et al. (2021) derived the explicit expressions of expected number of failures, expected duration of testing and Fisher information matrix for the unknown parameters of the underlying lifetime model. Then, using these quantities, the Bayesian optimal life testing plans under Type II unified hybrid censoring scheme were computed.

All these mentioned works have been done under the assumption that the sample size was fixed. In some cases, it is almost impossible to have a permanent fixed sample size because some observations may be missing for various reasons. Consider a system which consists of some machines. Assume that each day a certain number of items are produced by this system depending on the number of available units which may vary day by day. The variation in the number of units is caused by some factors such as malfunctioning of a machine and insufficient number of operators. Also, let each machine should have some conditions, which with probability θ is satisfied, to be utilized by such a system. Thus, in a long term, we have a system with binomial random number of units on hand. Some other discrete distributions such as Poisson, geometric, negative binomial or more generally power series class of distributions can be considered for the random sample size. Some examples in this area have been presented in Srivastava (1973). The distribution of order statistics with random sample size was studied in Epstein (1949), Raghunandanan and Patil (1972), Consul (1984), Gupta and Gupta (1984) and Rohatgi (1987).

The problem of predicting the range in a future sample, based on ranges in the earlier samples was discussed in Lingappaiah (1986), when the sample size was a random variable. Soliman (2000) studied the problem of predicting the future ordered observations in a sample from a Pareto distribution where the first r ordered observations have been observed and the sample size was a random variable having a Poisson or binomial distribution. The problem of predicting future generalized order statistics was discussed in Basiri and Ahmadi (2015), when the future sample size was a random variable. Since k-records and progressively Type II censored order statistics are contained in the model of generalized order statistics, the corresponding results for them were deduced as special cases. Assuming the information sample size as a random variable, Ahmadi et al. (2016) obtained the optimal parameter for the distribution of the information sample size such that the point predictor of a future order statistic had minimum mean squared prediction error and the total cost of experiment was bounded. Similar problem for predicting future order statistics based on upper and lower records was studied by Basiri and Pakzad (2018). Basiri and Asgharzadeh (2021) considered the progressively Type II censoring and the proportional hazard rate models. Then they introduced a cost function and the optimal sample size was found by minimizing the cost function, when the sample size was a random variable. Some known candidates for the random sample size, namely the degenerate, binomial, and Poisson distributions were considered in that paper. Parameteric and nonparametric prediction intervals based on generalized order statistics were respectively obtained by Barakat et al. (2018) and Barakat et al. (2021), when informative sample size was fixed as well as a random variable. Barakat and Newer (2022) constructed prediction intervals for future observations from the exponential and Pareto distributions in the context of ordered ranked set sampling. Two cases were assumed in their paper. The first case, when the sample size was a fixed value and the second case when the sample size was a positive integer-valued random variable. Basiri (2022) considered a parallel system with a random number of units. The number of units followed a power series distribution which includes several distributions such as geometric, logarithmic and zero-truncated Poisson distributions. Then by minimizing the expected cost and the reliability of this system, the optimal parameter for the distribution of sample size was obtained.

In the present article, the aim is to determine the optimal sample size in Type I hybrid censoring so that the total cost of the experiment does not exceed a pre-fixed value, C^*. Here, the sample size is assumed as a truncated binomial random variable. So, the optimal success probability of the binomial distribution is the aim of this paper.

The structure of the article is as follows. First, a cost function is introduced in Type I hybrid censoring model. Here, the sample size is a truncated binomial random variable. So, the success probability of binomial distribution is determined in such a way that the cost function is less than a pre-fixed value C^*. Numerical calculations and a simulation study are also provided. Finally, the conclusion is expressed.

2 Main results

Suppose a random sample $\tilde{Y} = (Y_1, \ldots, Y_N)$ consisting of N units from the Rayleigh distribution with pdf and cdf given in (1) and (2), has been tested. Also, let $Y_{1:N}, \ldots, Y_{N:N}$ be the corresponding order statistics. Assuming $N = n$, the pdf and cdf of $Y_{i:N}$, $1 \le i \le N$, respectively are given by (David and Nagaraja, 2003)

$$g_{Y_{i:N}|N=n}(y) = \frac{n!}{(i-1)!(n-i)!} g_\sigma(y) \left(G_\sigma(y)\right)^{i-1} \left(\bar{G}_\sigma(y)\right)^{n-i}, \tag{3}$$

$$G_{Y_{i:N}|N=n}(y) = \sum_{j=i}^{n} \binom{n}{j} \left(G_\sigma(y)\right)^j \left(\bar{G}_\sigma(y)\right)^{n-j}, \tag{4}$$

where $\bar{G}_\sigma(y) = 1 - G_\sigma(y)$. Independently, let N be a non-negative integer-valued random variable from the truncated binomial distribution at point m, say $TB(M, \theta; m)$, i.e.,

$$P(N = n) = \rho(m, M, \theta) \binom{M}{n} \theta^n (1-\theta)^{M-n}, \quad n = m, m+1, \ldots, M, \tag{5}$$

where $0 \le \theta \le 1$ is the success probability and

$$\rho(m, M, \theta) = \frac{1}{\sum_{n=m}^{M} \binom{M}{n} \theta^n (1-\theta)^{M-n}}. \tag{6}$$

In this case, we have

$$\begin{aligned}
\mathbb{E}_N(N) &= \rho(m, M, \theta) \sum_{n=m}^{M} n \binom{M}{n} \theta^n (1-\theta)^{M-n} \\
&= M\theta \rho(m, M, \theta) \sum_{n=m-1}^{M-1} \binom{M-1}{n} \theta^n (1-\theta)^{M-1-n} \\
&= M\theta \frac{\rho(m, M, \theta)}{\rho(m-1, M-1, \theta)}.
\end{aligned}$$

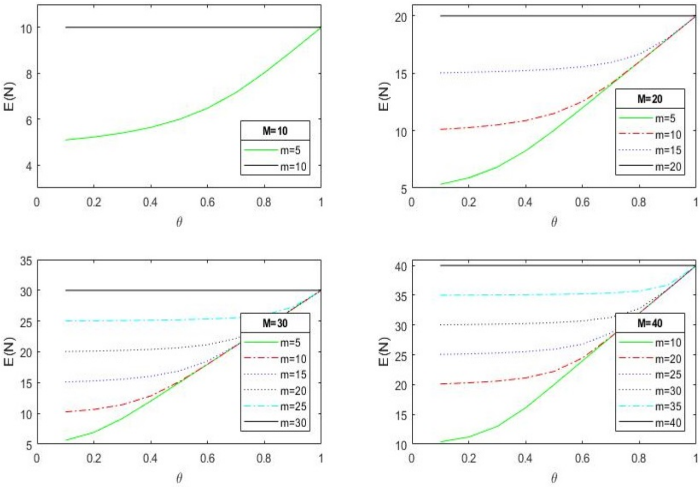

Figure 1: The values of $\mathbb{E}_N(N)$ with respect to θ for different choices of M and m.

Figure 1 shows graphical representations of $\mathbb{E}_N(N)$ with respect to θ for different choices of M and m. From Figure 1, we see that $\mathbb{E}_N(N)$ is increasing in θ, when m and M are kept fixed.

In the context of Type I hybrid censoring, we denote Δ and $\tau = \min(Y_{m:N}, Y_0)$ as the number of failures and the duration of the test, respectively, in which m and Y_0 are pre-fixed values. Clearly, Δ and τ are both random variables.

In order to determine the optimal number of units, N, in Type I hybrid censoring, the following cost function is first considered.

$$\Psi(\mathcal{C}) = C_0 + C_N \mathbb{E}_N(N) + C_f \mathbb{E}_N\left(\mathbb{E}(\Delta|N=n)\right) + C_t \mathbb{E}_N\left(\mathbb{E}(\tau|N=n)\right),$$

where $\mathcal{C} = (N, m, Y_0)$ is the censoring scheme and C_0, C_N, C_f and C_t are the sampling set-up cost or any other related costs involved in sampling, the cost per unit, the cost per failed item and the cost per unit of duration of life testing, respectively.

In this paper N is a random variable but $\mathbb{E}_N(N)$ is always fixed which is an increasing function of θ (see Figure 1). So, the optimal value for θ, say θ_{opt}, is the aim of paper.

On the other hand, for Type I hybrid censoring since

$$\min(Y_{m:N}, Y_0) = \begin{cases} Y_0 & Y_0 < Y_{m:N}, \\ Y_{m:N} & Y_0 \geq Y_{m:N}, \end{cases}$$

the data are $(Y_{1:N}, Y_{2:N}, \ldots, Y_{\Delta:N}, \Delta)$ and we obtain (Bhattacharya et al., 2014)

$$\begin{aligned} P(\Delta = j) &= \binom{N}{j} \left(G_\sigma(Y_0)\right)^j \left(\bar{G}_\sigma(Y_0)\right)^{N-j}, \quad j = 0, 1, \ldots, m-1, \\ P(\Delta = m) &= \sum_{j=m}^{N} \binom{N}{j} \left(G_\sigma(Y_0)\right)^j \left(\bar{G}_\sigma(Y_0)\right)^{N-j}. \end{aligned}$$

So, we can write (Bhattacharya et al., 2014)

$$\begin{aligned}\mathbb{E}(\Delta|N=n) &= \sum_{j=0}^{m-1} j\binom{n}{j}\left(G_\sigma(Y_0)\right)^j\left(\bar{G}_\sigma(Y_0)\right)^{n-j}\\ &\quad +m\sum_{j=m}^{n}\binom{n}{j}\left(G_\sigma(Y_0)\right)^j\left(\bar{G}_\sigma(Y_0)\right)^{n-j}. \end{aligned} \tag{7}$$

From (2) and (7), and using the binomial expansion for $(G_\sigma(Y_0))^j = \left(1-\bar{G}_\sigma(Y_0)\right)^j$, we find that

$$\begin{aligned}\mathbb{E}(\Delta|N=n) &= \sum_{j=0}^{m-1}\sum_{k=0}^{j} j\binom{n}{j}\binom{j}{k}(-1)^k\left(\bar{G}_\sigma(Y_0)\right)^{n-j+k}\\ &\quad +m\sum_{j=m}^{n}\sum_{k=0}^{j}\binom{n}{j}\binom{j}{k}(-1)^k\left(\bar{G}_\sigma(Y_0)\right)^{n-j+k}\\ &= \sum_{j=0}^{m-1}\sum_{k=0}^{j} j\binom{n}{j}\binom{j}{k}(-1)^k\exp\left(-\frac{(n-j+k)Y_0^2}{2\sigma^2}\right)\\ &\quad +m\sum_{j=m}^{n}\sum_{k=0}^{j}\binom{n}{j}\binom{j}{k}(-1)^k\exp\left(-\frac{(n-j+k)Y_0^2}{2\sigma^2}\right). \end{aligned} \tag{8}$$

On the other hand, we can write

$$\begin{aligned}\mathbb{E}(\tau|N=n) &= E(\min(Y_{m:N},Y_0)|N=n)\\ &= Y_0P(Y_{m:N}>Y_0|N=n)+\int_0^{Y_0} y g_{Y_{m:N}|N=n}(y)dy\\ &= Y_0\bar{G}_{Y_{m:N}|N=n}(Y_0)+\int_0^{Y_0} y g_{Y_{m:N}|N=n}(y)dy, \end{aligned} \tag{9}$$

where $g_{Y_{m:N}|N=n}(y)$ and $\bar{G}_{Y_{m:N}|N=n}(y) = 1-G_{Y_{m:N}|N=n}(y)$ are defined in (3) and (4), respectively. Integrating by parts leads to

$$\int_0^{Y_0} y g_{Y_{m:N}|N=n}(y)dy = Y_0 G_{Y_{m:N}|N=n}(Y_0) - \int_0^{Y_0} G_{Y_{m:N}|N=n}(y)dy. \tag{10}$$

From (9) and (10), we conclude that

$$\begin{aligned}\mathbb{E}(\tau|N=n) &= Y_0\bar{G}_{Y_{m:N}|N=n}(Y_0)+Y_0G_{Y_{m:N}|N=n}(Y_0)-\int_0^{Y_0} G_{Y_{m:N}|N=n}(y)dy\\ &= Y_0-\int_0^{Y_0} G_{Y_{m:N}|N=n}(y)dy. \end{aligned} \tag{11}$$

From (4), we have

$$\int_0^{Y_0} G_{Y_{m:N}|N=n}(y)dy = \sum_{j=m}^{n}\binom{n}{j}\int_0^{Y_0}\left(G_\sigma(y)\right)^j\left(\bar{G}_\sigma(y)\right)^{n-j}dy$$

$$= \sum_{j=m}^{n}\sum_{k=0}^{j}\binom{n}{j}\binom{j}{k}(-1)^k\int_0^{Y_0}\left(\bar{G}_\sigma(y)\right)^{n-j+k}dy, \quad (12)$$

where the last equality is obtained by using the binomial expansion for $(G_\sigma(y))^j = \left(1-\bar{G}_\sigma(y)\right)^j$. From (2) we get

$$\begin{aligned}\int_0^{Y_0}\left(\bar{G}_\sigma(y)\right)^{n-j+k}dy &= \int_0^{Y_0}\exp\left(-\frac{(n-j+k)y^2}{2\sigma^2}\right)dy\\ &= \frac{\sqrt{2\pi}\sigma}{\sqrt{n-j+k}}\int_0^{\frac{\sqrt{n-j+k}}{\sigma}Y_0}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{2}\right)dz\\ &= \frac{\sqrt{2\pi}\sigma}{\sqrt{n-j+k}}\left\{\Phi\left(\frac{\sqrt{n-j+k}}{\sigma}Y_0\right)-\frac{1}{2}\right\}, \quad (13)\end{aligned}$$

where $\Phi(\cdot)$ represents the cdf of the standard normal distribution, $N(0,1)$.

Now, from (11), (12) and (13), we find

$$\mathbb{E}(\tau|N=n) = Y_0 - \varphi_\sigma(m,n,Y_0), \quad (14)$$

in which

$$\varphi_\sigma(m,n,Y_0) = \sum_{j=m}^{n}\sum_{k=0}^{j}\binom{n}{j}\binom{j}{k}\frac{(-1)^k\sqrt{2\pi}\sigma}{\sqrt{n-j+k}}\left\{\Phi\left(\frac{\sqrt{n-j+k}}{\sigma}Y_0\right)-\frac{1}{2}\right\}.$$

Finally, (5) leads to

$$\mathbb{E}_N\left(\mathbb{E}(\Delta|N=n)\right) = \rho(m,M,\theta)\sum_{n=m}^{M}\mathbb{E}(\Delta|N=n)\binom{M}{n}\theta^n(1-\theta)^{M-n},$$

$$\mathbb{E}_N\left(\mathbb{E}(\tau|N=n)\right) = \rho(m,M,\theta)\sum_{n=m}^{M}\mathbb{E}(\tau|N=n)\binom{M}{n}\theta^n(1-\theta)^{M-n},$$

where $\rho(m,M,\theta)$, $\mathbb{E}(\Delta|N=n)$ and $\mathbb{E}(\tau|N=n)$ are defined in (6), (8) and (14), respectively.

In Table 1, the values of the cost function $\Psi(\mathcal{C})$ with different choices for Y_0, m, M and θ are reported, when $\sigma=1$, $C_0=10$, $C_N=15$, $C_f=10$, $C_t=20$. Figure 2 represents the values of $\Psi(\mathcal{C})$ with respect to m for different choices of M and θ, when $Y_0=0.5$, $\sigma=1$, $C_0=10$, $C_N=15$, $C_f=10$, $C_t=20$. From Table 1 and Figure 2 we observe that $\Psi(\mathcal{C})$ is increasing in Y_0, m, M and θ.

In the following, it is tried to find the optimal value for the parameter θ in $TB(M,\theta;m)$ in such a way that $\Psi(\mathcal{C}) \le C^*$, where C^* is a pre-fixed value. In Table 2 the values of θ_{opt} with different choices for m, M, C^* and Y_0 are reported which the condition $\Psi(\mathcal{C}) \le C^*$ is satisfied, when $\sigma=1$, $C_0=10$, $C_N=15$, $C_f=10$, $C_t=20$. The results in Table 2 show that, θ_{opt} may not be available based on the condition $\Psi(\mathcal{C}) \le C^*$. These situations are indicated by a dash $(-)$ in this table. Also, for some cases for all θ in $[0,1]$ the condition $\Psi(\mathcal{C}) \le C^*$ is satisfied. Figures 3 and 4 represent the values of $\Psi(\mathcal{C})$ with respect to θ for different choices of Y_0, m and M, when $C^*=400$, $\sigma=1$, $C_0=10$, $C_N=15$, $C_f=10$, $C_t=20$.

Table 1: The values of $\Psi(\mathcal{C})$ for different choices for Y_0, m, M and θ, when $\sigma = 1$, $C_0 = 10$, $C_N = 15$, $C_f = 10$ and $C_t = 20$.

Y_0	M	$m/\ \theta$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.5	10	5	103.95	107.60	112.00	117.49	124.50	133.62	145.40	159.78	176.31
		10	181.75	181.75	181.75	181.75	181.75	181.75	181.75	181.75	181.75
	20	5	110.11	122.03	139.03	162.85	192.28	224.10	256.38	288.66	320.96
		10	183.91	187.15	192.09	199.78	211.95	230.88	257.50	288.83	321.21
		15	263.74	265.11	266.86	269.17	272.35	277.05	284.59	297.94	321.82
		20	343.50	343.50	343.50	343.50	343.50	343.50	343.50	343.50	343.50
	30	5	116.68	140.71	178.06	224.24	272.51	320.86	369.10	417.17	465.05
		10	186.58	195.09	210.03	235.72	274.52	321.25	369.68	418.20	466.73
		15	265.99	270.34	276.40	285.71	301.49	329.02	370.39	418.21	466.73
		20	344.66	346.26	348.55	352.02	357.62	367.41	385.98	420.06	466.74
		25	425.04	425.86	426.92	428.33	430.30	433.24	438.13	447.68	470.46
		30	505.26	505.26	505.26	505.26	505.26	505.26	505.26	505.26	505.26
	40	5	124.05	164.92	224.37	288.63	353.01	417.11	480.86	543.95	600.06
		10	189.76	206.08	238.04	290.45	353.55	418.20	482.90	547.20	603.44
		15	268.32	276.30	289.35	313.76	357.89	418.39	482.87	546.62	610.61
		20	345.95	349.73	355.85	366.50	386.71	425.15	483.24	547.06	597.90
		25	426.38	428.95	432.43	437.51	445.77	461.19	492.98	549.28	631.10
		30	506.06	507.13	508.61	510.79	514.20	520.12	531.98	560.53	643.98
		35	589.40	590.01	590.85	592.06	593.89	596.84	602.17	614.38	677.54
		40	684.02	684.02	684.02	684.02	684.02	684.02	684.02	684.02	684.02
1	10	5	129.07	134.24	140.41	147.94	157.30	169.05	183.55	200.46	219.11
		10	219.35	219.35	219.35	219.35	219.35	219.35	219.35	219.35	219.35
	20	5	137.73	153.77	175.11	203.02	235.56	269.06	301.62	333.06	363.71
		10	222.74	227.53	234.49	244.84	260.55	284.02	315.90	352.52	389.77
		15	315.58	317.52	319.97	323.20	327.65	334.12	344.33	361.77	391.32
		20	408.69	408.69	408.69	408.69	408.69	408.69	408.69	408.69	408.69
	30	5	146.67	176.99	219.75	268.79	317.03	363.32	408.39	452.92	497.29
		10	226.68	238.52	257.94	289.59	335.55	389.33	443.45	495.65	545.59
		15	318.75	324.82	333.13	345.51	365.62	399.27	448.33	504.25	560.59
		20	410.52	412.97	416.36	421.30	428.96	441.84	465.19	506.32	561.24
		25	504.30	505.47	506.96	508.95	511.71	515.83	522.58	535.45	564.63
		30	598.04	598.04	598.04	598.04	598.04	598.04	598.04	598.04	598.04
	40	5	156.31	205.00	268.72	332.45	393.34	452.89	512.11	571.51	634.77
		10	231.20	252.85	292.27	353.86	425.26	495.04	561.10	624.04	686.83
		15	322.01	332.97	350.15	380.61	433.45	504.34	578.91	652.00	725.45
		20	412.50	418.03	426.50	440.51	465.81	512.10	580.52	655.73	731.08
		25	506.20	509.82	514.67	521.63	532.62	552.33	591.11	656.19	728.09
		30	599.30	600.95	603.18	606.35	611.16	619.19	634.40	667.51	731.36
		35	693.38	694.21	695.28	696.70	698.68	701.66	706.60	716.33	741.77
		40	787.59	787.59	787.59	787.59	787.59	787.59	787.59	787.59	787.59
2	10	5	173.76	180.14	187.58	196.30	206.53	218.38	231.61	245.42	259.66
		10	286.47	286.47	286.47	286.47	286.47	286.47	286.47	286.47	286.47
	20	5	184.28	202.18	222.86	246.66	272.77	300.11	328.16	356.68	385.59
		10	291.77	298.77	308.19	321.03	338.57	361.64	388.88	416.95	444.78
		15	407.03	409.91	413.55	418.28	424.70	433.83	447.56	469.14	500.32
		20	522.93	522.93	522.93	522.93	522.93	522.93	522.93	522.93	522.93
	30	5	194.46	224.32	260.13	300.35	342.55	385.67	429.29	473.23	517.38
		10	297.53	313.13	335.19	365.60	403.59	444.97	487.32	530.34	573.82
		15	411.73	420.56	432.13	448.08	471.12	504.03	544.93	587.41	629.98
		20	526.01	529.99	535.33	542.78	553.72	570.81	598.60	640.35	686.27
		25	642.62	644.42	646.73	649.79	654.03	660.27	670.34	688.76	726.72
		30	759.40	759.40	759.40	759.40	759.40	759.40	759.40	759.40	759.40
	40	5	204.61	248.05	300.50	356.93	414.77	473.28	532.17	591.01	634.66
		10	303.71	329.57	367.58	417.68	473.27	530.46	588.47	646.70	689.64
		15	416.50	431.84	453.39	485.71	532.28	587.51	644.46	701.95	745.38
		20	529.23	537.84	550.11	568.61	598.01	643.49	700.23	757.43	800.17
		25	645.55	651.11	658.45	668.68	684.02	709.15	751.96	810.73	849.97
		30	761.54	764.28	767.92	772.96	780.38	792.20	813.20	854.21	913.23
		35	878.68	879.98	881.65	883.87	886.97	891.58	899.17	913.77	949.69
		40	995.87	995.87	995.87	995.87	995.87	995.87	995.87	995.87	995.87

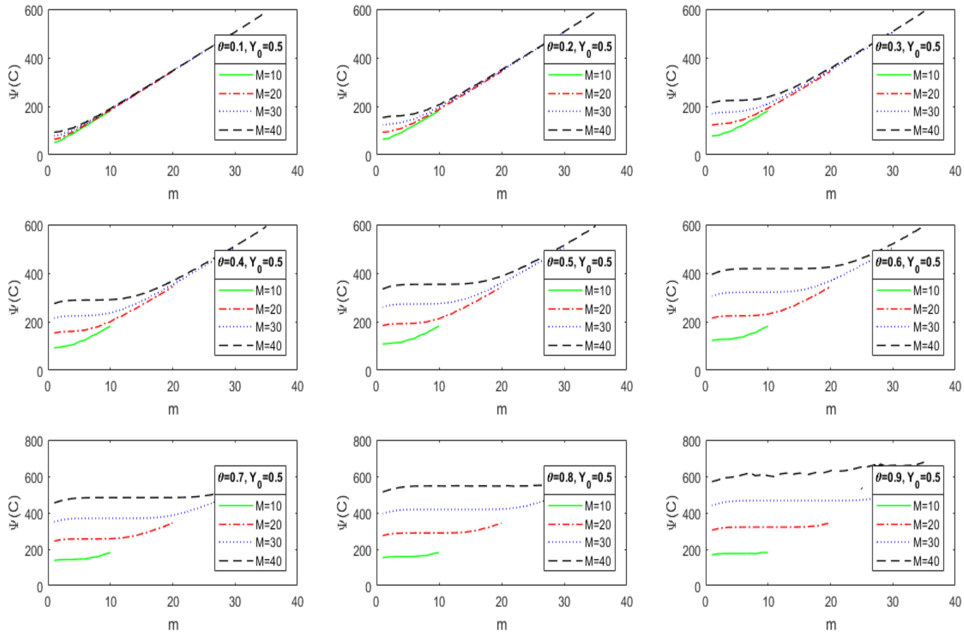

Figure 2: The values of $\Psi(\mathcal{C})$ with respect to m for different choices of M and θ, when $Y_0 = 0.5$, $\sigma = 1$, $C_0 = 10$, $C_N = 15$, $C_f = 10$, $C_t = 20$.

Table 2: The values of θ_{opt} for different choices of Y_0, m, M, when $\sigma = 1$, $C_0 = 10$, $C_N = 15$, $C_f = 10$, $C_t = 20$ and $C^* = 400,\ 500$.

	C^*	500			400		
M	m	$Y_0 = 0.5$	$Y_0 = 1$	$Y_0 = 2$	$Y_0 = 0.5$	$Y_0 = 1$	$Y_0 = 2$
10	5	[0, 1]	[0, 1]	[0, 1]	[0, 1]	[0, 1]	[0, 1]
	10	[0, 1]	[0, 1]	[0, 1]	[0, 1]	[0, 1]	[0, 1]
20	5	[0, 1]	[0, 1]	[0, 1]	[0, 1]	[0, 1]	[0, 0.9417]
	10	[0, 1]	[0, 1]	[0, 1]	[0, 1]	[0, 0.9243]	[0, 0.7396]
	15	[0, 1]	[0, 1]	[0, 0.899]	[0, 1]	[0, 0.9209]	–
	20	[0, 1]	[0, 1]	–	[0, 1]	–	–
30	5	[0, 0.9670]	[0, 0.9055]	[0, 0.8606]	[0, 0.7643]	[0, 0.6814]	[0, 0.6329]
	10	[0, 0.9627]	[0, 0.8087]	[0, 0.7295]	[0, 0.7625]	[0, 0.6197]	[0, 0.4906]
	15	[0, 0.9626]	[0, 0.7924]	[0, 0.5878]	[0, 0.7619]	[0, 0.6015]	–
	20	[0, 0.9621]	[0, 0.7846]	–	[0, 0.7411]	–	–
	25	[0, 0.9599]	–	–	–	–	–
	30	–	–	–	–	–	–
40	5	[0, 0.7303]	[0, 0.6796]	[0, 0.6454]	[0, 0.5733]	[0, 0.5112]	[0, 0.4745]
	10	[0, 0.7266]	[0, 0.6075]	[0, 0.5467]	[0, 0.5718]	[0, 0.4646]	[0, 0.3647]
	15	[0, 0.7263]	[0, 0.5939]	[0, 0.4307]	[0, 0.5696]	[0, 0.4367]	–
	20	[0, 0.7261]	[0, 0.5739]	–	[0, 0.5346]	–	–
	25	[0, 0.7125]	–	–	–	–	–
	30	–	–	–	–	–	–
	35	–	–	–	–	–	–
	40	–	–	–	–	–	–

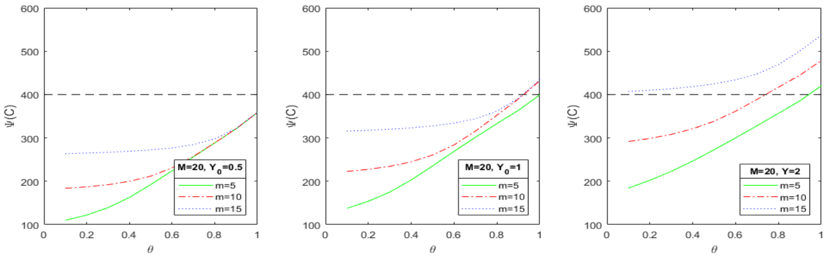

Figure 3: The values of $\Psi(\mathcal{C})$ with respect to θ for different choices of Y_0 and m, when $C^* = 400$, $M = 20$, $\sigma = 1$, $C_0 = 10$, $C_N = 15$, $C_f = 10$, $C_t = 20$.

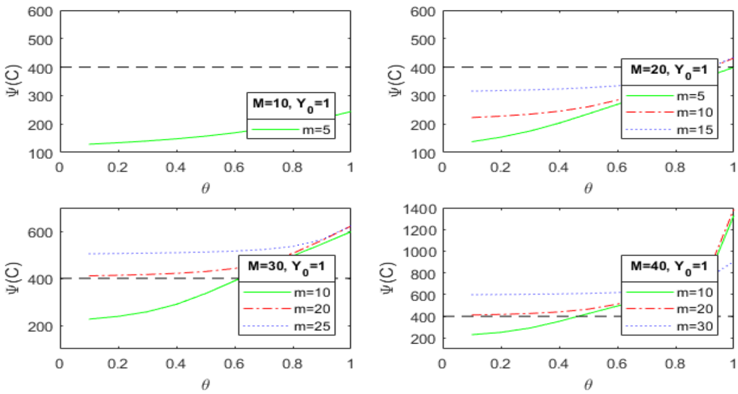

Figure 4: The values of $\Psi(\mathcal{C})$ with respect to θ for different choices of m and M, when $C^* = 400$, $Y_0 = 1$, $\sigma = 1$, $C_0 = 10$, $C_N = 15$, $C_f = 10$, $C_t = 20$.

3 Monte Carlo Simulation Study

In this section, we run a Monte Carlo simulation to assess the performances of the results in Sections 2. The following algorithm has been applied for this purpose:

Algorithm 1: Suppose m, M, Y_0, σ, C^*, C_0, C_N, C_f and C_t are all given. Then:

1. Choose θ_{opt} such that the condition $\Psi(\mathcal{C}) \leq C^*$ is satisfied.
2. Generate N from the distribution $TB(M, \theta_{opt}; m)$, say N_{opt}.
3. Generate N_{opt} iid random variables $Y_1, \ldots, Y_{N_{opt}}$ from the one-parameter Rayleigh distribution with parameter σ. Then sort them as $Y_{(1)} \leq \ldots \leq Y_{(N_{opt})}$.
4. Set $T = Y_0$ and $D = \sum_{j=1}^{m} I(Y_{(j)} \leq Y_0)$ if $Y_0 < Y_{(m)}$, else set $T = Y_{(m)}$ and $D = m$, where $I(\cdot)$ denotes the indicator function.
5. Repeat the Steps 2-4 for $B = 10000$ number of times and let $N_{opt}(i)$, $D(i)$ and $T(i)$ be the obtained results from Steps 2 and 4 in the i-th iteration, $i = 1, \ldots, B$.
6. Then, calculate the expected cost function $(E\Psi(\mathcal{C}))$ by using

$$E\Psi(\mathcal{C}) = C_0 + \frac{C_N}{B}\sum_{i=1}^{B} N_{opt}(i) + \frac{C_f}{B}\sum_{i=1}^{B} D(i) + \frac{C_t}{B}\sum_{i=1}^{B} T(i).$$

Table 3: The values of $\left(\theta_{opt},\ E\Psi(\mathcal{C})\right)$ with different choices for Y_0, m, M, when $\sigma = 1$, $C_0 = 10$, $C_N = 15$, $C_f = 10$, $C_t = 20$ and $C^* = 500$.

M	m	$Y_0 = 0.5$	$Y_0 = 1$	$Y_0 = 2$
10	5	(0.5, 110.8075)	(0.5, 136.6488)	(0.5, 173.4826)
		(1, 181.6954)	(1, 216.0941)	(1, 232.0981)
	10	(0.5, 181.8130)	(0.5, 219.3019)	(0.5, 285.444)
		(1, 181.8430)	(1, 219.3950)	(1, 285.5316)
20	5	(0.5, 181.2999)	(0.5, 214.7497)	(0.5, 232.7446)
		(1, 342.8296)	(1, 373.6571)	(1, 374.5493)
	10	(0.5, 196.112)	(0.5, 236.2446)	(0.5, 301.1152)
		(1, 343.5900)	(1, 406.1803)	(1, 432.8814)
	15	(0.5, 262.6975)	(0.5, 314.1810)	(0.4495, 404.3159)
		(1, 343.3200)	(1, 408.5508)	(0.899, 460.1978)
	20	(0.5, 343.4920)	(0.5, 408.6100)	–
		(1, 343.6240)	(1, 408.6630)	–
30	5	(0.4835, 254.4559)	(0.4527, 274.3704)	(0.4303, 273.3157)
		(0.967, 486.7389)	(0.9055, 479.7269)	(0.8606, 460.2932)
	10	(0.4813, 254.5250)	(0.4043, 264.6405)	(0.3647, 311.2633)
		(0.9627, 486.9158)	(0.8087, 480.1950)	(0.7295, 460.0494)
	15	(0.4813, 275.494)	(0.3962, 316.8379)	(0.2939, 404.4795)
		(0.9626, 487.4865)	(0.7924, 479.8193)	(0.5878, 455.4812)
	20	(0.4810, 343.6705)	(0.3923, 408.8500)	–
		(0.9621, 486.776)	(0.7846, 477.228)	–
	25	(0.4799, 424.4530)	–	–
		(0.9599, 485.6565)	–	–
40	5	(0.3651, 256.9612)	(0.3398, 274.6762)	(0.3227, 273.3488)
		(0.7303, 489.8751)	(0.6796, 480.3597)	(0.6454, 460.5287)
	10	(0.3633, 256.2685)	(0.3037, 266.9545)	(0.2733, 312.5757)
		(0.7266, 490.1829)	(0.6075, 479.607)	(0.5467, 459.5492)
	15	(0.3631, 278.2255)	(0.2969, 318.0100)	(0.2153, 404.3038)
		(0.7263, 489.5140)	(0.5939, 478.5786)	(0.4307, 451.3447)
	20	(0.3630, 344.0930)	(0.2869, 408.9100)	–
		(0.7261, 489.0280)	(0.5739, 470.0419)	–
	25	(0.3562, 424.3140)	–	–
		(0.7125, 483.3320)	–	–

Based on Algorithm 1 and the results in Table 2, we have computed the values of $E\Psi(\mathcal{C})$ with different choices for Y_0, m, M, when $\sigma = 1$, $C_0 = 10$, $C_N = 15$, $C_f = 10$, $C_t = 20$ and $C^* = 500$. All the obtained results are reported in Table 3. For the first step of Algorithm 1, we have selected the values of θ_{opt} from Table 2. Two points as the middle and the upper limit of the presented intervals are considered for computing the results. We mention that the results reported in Table 3 are based on 10000 Monte Carlo simulations. As expected, from Table 3, it is observed that in all cases $E\Psi(\mathcal{C}) \leq C^* = 500$. It is further observed that as expected, as θ_{opt} increases the values of $E\Psi(\mathcal{C})$ increase.

4 Conclusion

Combining Type I and Type II censoring methods results in hybrid censoring scheme. Determining the best number of units in a hybrid censoring scheme is an interesting issue. Practically, it is preferable for the experimenter to acquire the sample size that reduces the cost. In this article, the sample size is considered as a random variable of a truncated binomial distribution, and the optimal parameter of this distribution is determined in such a way that the cost function does not exceed a pre-determined value.

Acknowledgements

The authors express their sincere thanks to the editor-in-chief and the referees for their constructive comments and suggestions on the original version of this manuscript, which led to this improved version.

References

Ahmadi, J., Basiri, E. and MirMostafaee, S.M.T.K. (2016). Optimal random sample size based on Bayesian prediction of exponential lifetime and application to real data. *Journal of the Korean Statistical Society*, **45**(2):221–237.

Asgharzadeh, A. and Azizpour, M. (2016). Bayesian inference for Rayleigh distribution under hybrid censoring. *International Journal of System Assurance Engineering and Management*, **7**:239–249.

Barakat, H.M., El-Adll, M.E. and Aly, A.E. (2021). Two-sample nonparametric prediction intervals based on random number of generalized order statistics. *Communications in Statistics-Theory and Methods*, **50**(19):4571–4586.

Barakat, H.M. and Newer, H.A. (2022). Exact prediction intervals for future exponential and Pareto lifetimes based on ordered ranked set sampling of non-random and random size. *Statistical Papers*, **63**:1801–1827.

Barakat, H.M., Nigm, E.M., El-Adll, M.E. and Yusuf, M. (2018). Prediction of future generalized order statistics based on exponential distribution with random sample size. *Statistical Papers*, **59**:605–631.

Basiri, E. (2017). Optimal number of failures in Type II censoring for Rayleigh distribution. *Journal of Applied Research on Industrial Engineering*, **4**(1):67–74.

Basiri, E. (2022). Bi-objective optimization problem of a parallel system with random number of units. *Journal of Statistical Modelling: Theory and Applications*, **3**(1):31–49.

Basiri, E. and Ahmadi, J. (2015). Prediction intervals for generalized-order statistics with random sample size. *Journal of Statistical Computation and Simulation*, **85**(9):1725–1741.

Basiri, E. and Asgharzadeh, A. (2021). Optimal random sample size in progressively Type-II censoring based on cost constraint for the proportional hazards family. *Journal of Statistical Computation and Simulation*, **91**(15):3154–3169.

Basiri, E. and Pakzad, A. (2018). Optimal non-parametric prediction intervals for order statistics with random sample size. *Iranian Journal of Management Studies*, **11**(2):307–322.

Bhattacharya, R., Pradhan, B. and Dewanji, A. (2014). Optimum life testing plans in presence of hybrid censoring: a cost function approach. *Applied Stochastic Models in Business and Industry*, **30**(5):519–528.

Burkschat, M., Cramer, E. and Kamps, U. (2006). On optimal schemes in progressive censoring. *Statistics & Probability Letters*, **76**(10):1032–1036.

Burkschat, M., Cramer, E. and Kamps, U. (2007). Optimality criteria and optimal schemes in progressive censoring. *Communications in Statistics-Theory and Methods*, **36**(7):1419–1431.

Burkschat, M. (2008). On optimal linear equivariant estimation under progressive censoring. *Statistics*, **42**(5):383–392.

Childs, A., Chandrasekar, B., Balakrishnan, N. and Kundu, D. (2003). Exact likelihood inference based on Type-I and Type-II hybrid censored samples from the exponential distribution. *Annals of the Institute of Statistical Mathematics*, **55**(2):319–330.

Consul, P.C. (1984). On the distributions of order statistics for a random sample size. *Statistica neerlandica*, **38**(4):249–256.

David, H.A. and Nagaraja, H.N. (2003). *Order Statistics*. 3rd ed., New Jersey: John Wiley & Sons.

Ebrahimi, N. (1988). Determining the sample size for a hybrid life test based on the cost function. *Naval Research Logistics*, **35**(1):63–72.

Epstein, B. (1949). A modified extreme value problem. *The Annals of Mathematical Statistics*, **20**(1):99–103.

Epstein, B. (1954). Truncated life test in the exponential case. *The Annals of Mathematical Statistics*, **25**:555–564.

Gupta, D. and Gupta, R.C. (1984). On the distribution of order statistics for a random sample size. *Statistica Neerlandica*, **38**(1):13–19.

Lingappaiah, G.S. (1986). Bayes prediction in exponential life-testing when sample size is a random variable. *IEEE Transactions on Reliability*, **35**(1):106–110.

MirMostafaee, S.M.T.K., Amini, M. and Asgharzadeh, A. (2017). Bayesian prediction of minimal repair times of a series system based on hybrid censored sample of components' lifetimes under Rayleigh distribution. *Communications in Statistics-Theory and Methods*, **46**(4):1788–1806.

Mishra, N. (2020). Optimal one-step censoring schemes under entropy criterion. *Communications in Statistics-Simulation and Computation*, **49**(8):2068–2081.

Prajapat, K., Koley, A., Mitra, S. and Kundu, D. (2021). An optimal Bayesian sampling plan for two-parameter exponential distribution under Type-I hybrid censoring. *Sankhya A*, **85**:512–539

Raghunandanan, K. and Patil, S.A. (1972). On order statistics for random sample size. *Statistica Neerlandica*, **26**(4):121–126.

Rohatgi, V.K. (1987). Distribution of order statistics with random sample size. *Communications in Statistics-Theory and Methods*, **16**(12):3739–3743.

Sen, T., Bhattacharya, R., Pradhan, B. and Tripathi, Y.M. (2021). Statistical inference and Bayesian optimal life-testing plans under Type II unified hybrid censoring scheme. *Quality and Reliability Engineering International*, **37**(1):78–89.

Soliman, A.A. (2000). Bayes prediction in a Pareto lifetime model with random sample size. *Journal of the Royal Statistical Society: Series D (The Statistician)*, **49**(1):51–62.

Srivastava, R.C. (1973). Estimation of probability density function based on random number of observations with applications. *International Statistical Review/Revue Internationale de Statistique*, **41**(1):77–86.

Volterman, W., Davies, K.F. and Balakrishnan, N. (2012). Pitman closeness as a criterion for the determination of the optimal progressive censoring scheme. *Statistical Methodology*, **9**(6):563–572.

Wu, S.J., Huang, S.R. and Wang, J.H. (2023). Determination of warranty length for one-shot devices with Rayleigh lifetime distribution. *Communications in Statistics-Theory and Methods*, **52**(5):1400–1416.