**Relative Loss of Information and Efficiency of Censored Life Testing Experiments based on Sufficient Statistics**

**Hatim Solayman Migdadi**

Department of Mathematics, Faculty of Science and Information Technology, Jadara University, B.O. Box: 733, Irbid 21110, Jordan. E-mail:hmigdadi@jadara.edu.jo

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Abstract**: In this paper, sufficient statistics are proposed to measure the amount of information in censored and progressively censored data. Under specified life testing experiments, the sufficient statistics are obtained and used to identify the relative losses of information and the efficiency due to the censoring schemes. Comparisons between the different censoring mechanisms, the main affecting factors in the amount of information and the validity of sufficient statistics as measures of information are explored through both theoretical and numerical results.

**Keywords**: Type I and Type II censored data, Progressively censored data, Sufficient statistic, Relative loss of information, Relative efficiency.

**1-Introduction**

 In reliability and survival studies. Life testing experiments are performed to provide the failure times data required for any statistical inference about the lifetime model. Under the normal used conditions most units need a long time for failure. Thus having a complete data set is costly and time consuming. To overcome this problem, the life testing experiments are frequently incorporate censoring schemes with a relative loss of data that affect the efficiency for any statistical modeling.

The most censoring mechanisms are type I (time) censoring, where the life testing experiment will be terminated at a prescribed time, and type II (failure) censoring ,where the life testing experiment will be terminated upon the failure, where, is a pre-fixed) number of failures. Using both types of censoring, progressively censored data my obtained when at various stages of the experiment, some of the surviving units are eliminated from further observations.

 Statisticians have long endeavored to develop a precise notation for the information in the sampling data. Fisher information is basically considered by many authors with several censoring schemes as in.[1],[2],[3] . It is also involved in .[4] for the progressively type II censoring. Fisher information in the generalized order statistics is studied by [5]. Others of many related references are in([6]-[9]).

Despite the importance of fisher information measure, the differential entropy measure introduced by Shannon.[10] is utilized in the advanced information theory. Statistically it used by many authors as in [11],[12], [13] and [14]For continuous random variable distributed with probability distribution function , it is defined by

 (1)

For the negativity of Shannon information in some cases, Awad(1987) suggested a modification for Shannon entropy by

 , where (2)

This measure is used by [15] for the Pareto distribution with type II censoring and by [16] for obtaining the relative efficiency of the information in type I censored sample at time from the exponential distribution.

A sufficient statistic is a statistic that in a certain sense captures all the information about the parameter of interest in the sample data, and any conditional information beside the value of the sufficient statistics does not contain any more information.

The aim of this paper is to investigate sufficient statistics of the exponential scale parameter of the distribution with probability distribution and reliability functions given respectively by

 (3)

 (4)

as a measure of the information in the data obtained using different censoring schemes . The exponential distribution is considered because it plays an important basically role in survival and reliability analysis. Its widely used in the development theory of life testing because of its mathematical tractability as it belongs to both exponential and weibull family distributions . More details are found in [17]-[19].

Consider a complete random sample from the , then the likelihood function of is given by

 (5)

Applying the factorization theorem **,**  is a sufficient statistics of

We define the amount of information given the complete sample as

 (6)

For other censored life testing experimental data with sufficient statistic,the relative loss of information due to censoring is defined as

 (7)

and ,the relative efficiency of the censoring schemes as

 (8)

Consequently,

 (9)

The relative loss of information and efficiency are obtained for type I censored data in section 2 and for type I censored data in section 3,and for progressively type I and progressively censored data in the sections 4 and 5 respectively. In section 6 numerical results about the theoretical findings are illustrated. Conclusion of the overall paper is involved in section 7.

**2-Type 2 censored data**

Suppose that units are placed on a life testing experiment and the experiment is terminated when observing the -th failure, where is a pre fixed known constant. Hence, the only observable failure times data are and the remaining failure times are censored. Such type of data is referred to as type II censored data. Assume that the life times of the units are identically distributed follow the Exponential distribution, then the likelihood function of given the ordered failure times is

 (10)

By the factorization theorem ,sufficient statistics for based on the type II censored data is given by

 (11)

The amount of information utilized from the type II censored data

can be found using the following theorem.

**Theorem**: (Peter.2002.P129):If the lifetime variables are identically distributed by , then has a distribution.

This implies. Therefore, the relative loss of information due to type II censored data is

 (12)

and the efficiency of type II censored data is

 (13) Since, fisher information in the type II censored data is given by

 , this implies, the relative loss of fisher information due to type II censored data is given by

 (14)

**3-Type I censored data**

If the life testing experiment considered in section (2) is terminated at a pre-assigned time .Then we have type I censoring, with the lifetimes data:

 , the likelihood function of is given by

 (15)

Hence, sufficient statistics of is

 (16)

Since : ,clearly based on the sufficient statistics with same proportion of censored data , the amount of information using type I censored data is greater than the amount of information using type II censored data .

Since

This implies, as compared to type II censored data with failures, the a mount of information :

 Where

For type I censored data is a random variable distributed as

Hence , ,setting implies that an approximate value of can be obtained using the theory of order statistic as:

 (17)

This implies can be approximated by:

)

Hence ,the relative loss of information due to type I censored data is given by

 (18)

And the efficiency of type I censoring is given by

 (19)

4-Progressively Type I censored data

The progressively type I censored data are obtained from the life test as follows:

Assume that inspection times are predetermined and units are put on the life test for failure at time 0.At the inspection time, failure units in the interval are recorded, and of the surviving units are randomly removed. Continuing the test, at time , failure units in the interval are recorded,and of the are selected and randomly removed..finally ,at the censoring time , failure units in the interval are recorded and all the surviving units are removed and the test is stopped.

The likelihood function of based on the progressively type I censored data

 is given by

 (20)

Hence, sufficient statistics of is

To find ,we have to find ,

Setting: the conditional probability of failure in the interval given of survival at the time ,and assuming then

This implies:

 is distributed as

distributed as

If are proportion of the remaining surviving units, i.e. , and .By induction:

Substituting for in (21),the amount of information in progressively type I censored data is given by

(21)

Hence, the relative loss of information due to progressively type I censored data is

 (22)

And ,the relative efficiency of progressively type I censored data is

 (23)

5- Progressively type II censoring

Consider an experiment in which units are placed on a life testing experiment. At the time of the first failure, units are randomly removed from the remaining surviving units. Similarly, at the time of the second failure, units from the remaining units are randomly removed. The test continues until the failure at which time, all the remaining units are removed and the test stopped. Hence, ,the likelihood function of based on the progressive type II censored data

can be simplified as

 (24)

Where the constant is independent of ,by the factorization theorem ,sufficient statistics is given by

 (25)

Hence, either are predetermined constants ,or assumed to be binomially removals with probability ,i.e distributed as .to find ,assume are independent, from the theory of order statistics

 (26)

Since , ,this implies , the amount of information in the progressively type II censored data

Therefore, the relative loss of information due to progressively censored data is given by

(27)

and ,the efficiency of the progressively censored data is

 (28)

We notice that, if and ,then we have exactly type II censoring.

6-Numerical results

To illustrate theoretical finding in this paper, random samples of size from the distribution are generated with .Relative loss of information and efficiency of type I and type II censored data are presented and compared with the relative efficiency based on Awad Sup-Entropy given by

 (29)

 in Table 1.

Table 1. Relative loss of information and efficiency of type I and type II censored data

 Proportion

 of censoring

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

3 0.10 0.90 0.0910 0.9090 0.8366

 0.30 0.70 0.2914 0.7086 0.5769

 0.50 0.50 0.4922 0.5078 0.5063

5 0.10 0.90 0.0906 0.9094 0.8527

 0.30 0.70 0.2907 0.7093 0.6190

 0.50 0.50 0.4913 0.5087 0.5580

7 0.10 0.90 0.0903 0.9097 0.8605

 0.30 0.70 0.2902 0.7098 0.6372

 0.50 0.50 0.4909 0.5091 0.5767

It is clearly appears that:

(1) As proportions of censored data increases, the relative loss of information increases and the efficiency decreases using both types of censoring.

(2) For type I censored data with fixed proportion of censoring ,the relative efficiencies increases and the relative loss of information decreases as the values of the scale parameter increases, while type II censored data do not affected by the parameter values .

(3) Based on the sufficient statistics, the relative efficiencies of type I censored are greater than their corresponding values based on Awad sup entropy.

(4) Based on the sufficient statistics, the relative efficiencies of type I censored are greater than their corresponding values using type II censored data.

To measure the relative loss of information and efficiency of progressively type I censored data .equidistance inspection intervals are considered with number of inspections , and fixed probability of removals as presented in Table 2. It is clearly that:

Table 2. Relative loss of information and efficiency of progressively type I censored data

 Probability number of

 of removals inspections

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

3 0.05 4 0.3021 0.6979

 5 0.2987 0.7013

 6 0.2965 0.7035

 0.10 4 0.3201 0.6799

 5 0.3145 0.6855

 6 0.3008 0.6992

5 0.05 4 0.2996 0.7004

 5 0.2934 0.7066

 6 0.2871 0.7129

 0.10 4 0.3166 0.6834

 5 0.3067 0.6933

 6 0.2984 0.7016

7 0.05 4 0.2878 0.7122

 0.2806 0.7294 5

 6 0.2765 0.7235

 0.10 4 0.3001 0.6999

 5 0.2963 0.7037

 6 0.2892 0.7208

(1) The relative loss of information decreases and the relative efficiencies increases as the number of inspections increases.

(2) As the proportions of removals increasing from ,the relative losses of information increases and the relative efficiencies decreases.

(3) For fixed probability of removals and fixed number of inspections ,the relative loss of information decreases as values of the scale parameter increases .

To measure the relative loss of information and efficiency of progressively type II censored data .two, three and four stages are considered. For the two stages ,removals are assumed at the failure times with probability of removals in the first case and in the second case. For the three stages ,removals are assumed at the failure times with probability of removals in the first case and in the second case. For the four stages ,removals are assumed at the failure times with probability of removals:

 in the first case and in the second case. Otherwise probabilities of removals assumed to be identically 0.as it appears in Table 3. It is clearly that

Table 3. Relative loss of information and efficiency of progressively type I I censored data

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

 Stages case

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Two first 0.1214 0.8786

 second 0.1537 0.8463

Three first 0.1937 0.8073

 second 0.2185 0.7815

Four first 0.2285 0.7715

 second 0.2352 0.7648

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(1)while type I censored gave relatively better efficiencies than type II censored data ,progressively type II censored data have more relative efficiencies than progressively type I censored data with same proportion of removals. One explanation for this, is that in case of progressively type I, there exists extra loss of information about the exact failure times.

(2) As the probabilities of binomial removals increases, the relative loss of information increases and the relative efficiencies decreases.

(3) With the same of binomial removals, the relative losses of information increases and the relative efficiencies decreases as the numbers of stages of censoring increases and hence, the optimum censoring is the one stage ordinary type II censoring.

**7-Conclusion**

Based on sufficient statistics, the main factors that affects the relative losses of information and efficiencies of the censored data are, the proportion of censoring, the indexed parameter values and the length of the life testing experiments. For type II and progressively type II censoring, the amount of information is basically related with the proportion of censoring and probabilities of removal units at different stages of the life testing. While type I censoring provide more information than the ordinary type II. Progressively type II censoring manifests more relative efficiencies than progressively type I censoring. This is due to the specific loss of information associated with progressively type I in the exact failure times.

The overall theoretical and numerical found out results in this paper, ensures that sufficient statistics could be a probable measure for the amount of information in censored data. The developed methods can be generalized to other life time models using mathematical transformations and modern software tools .Future researches may highlights planning life testing experiments based on sufficient statistics with other censoring mechanisms.

**References**

[1] Zheng . G.Joseph.L. Gastwirth. On the fisher information in randomly censored data.2001. Statistics and Probability letters.52.421-426.

[2]Zheng.G. On the fisher information matrix in type II censored data from the exponetiated Exponential family.2002. Biometrical Journal.44.3.353-357.

[3] Zheng.G.Sangun.P.On the fisher information in multiply censored and progressively censored data .2004.Communications in Statistics-Theory and Methods.33.8.1821-1835.

[4]Dahmen.K.Burkshat.M.Cramer.E. A-and D-Optimal progressive type II censoring designs based on fisher information.2012.Journal of statistical computation and simulation. .2012.82.6.879-905.

[5] Park, S. (1996), Fisher information in order statistics. J. Amer. Statist.

Assoc., **91**, 385-390.

[6]Efron,B.andJohnstone,I.(1990),“Fisher information in terms of the hazard rate”,Annals of Statistics,vol.18,38-62.

[7]Luis.A.E. Meeker.Q.W.Fisher information matrices with censoring,truncation, and explanatory variables.1998. Statistica Sinica.8.221-227

[8]Balakrishnan,N.,Progressivecensoringmethodology:anappraisal,Test16(2007),211–296(withdiscussion

[9]Balakrishnan, N., Burkschat, M., Cramer, E. and Hofman, G. (2008). Fisher information based progressive censoring planes. COMSTA 4123, doi: 10.1016/j.csda.07.038

.[10] Shannon, C. E. (1948). A mathematical theory of communication. Bell System Technical Journal, 27, 379-423

[11] Hollander, M., Proshcan, F. and Sconing, J. (1987). Measuring information in right-censored models. Naval Rcs. Logist, 34, 669-681

[12]Ebrahimi N.,M.Habibullah and E.S.Soofi,.(1993).*The Role of information theory*

 *In Reliability analysis*. Elsevier Science Publishers B V.89-101

[13]Menendez ,. M.L. (2000) .*Shannon’s Entropy in Exponential Families*: Statistical

 Applications. Applied Mathematics Letters.13.37-42

[14]Mathai, . A.M. and Haubold .H.J.(2007).*Pathway model, Super statistics* ,.*Pathway model, Super statistics , Tsallis Statistics ,and a generalized measure of entropy*. Physica A 375.110- 122.

[15]Haj Ahmed, H. A. and Awad, A. M. (2010). Optimality criterion for progressive type-II right censoring based on Awad sup-entropy measures. Journal of Statistics, 16, 12-27

[16][Kittaneh.O.A.Deriving the efficiency function for type I censored sample from exponential distribution using sup-entropy.2012.Journal of Statistics.19.43-53

[17]Balakrishnan, N., Basu,A.P., 1995. The Exponential Distribution: Theory, Methods and Applications. Gordon and Beach Publishers, The Netherlands.

[18]Barlow.E , and Proschan, *F.Statistical Theory of Reliability and Life Testing: Probability Models*, Silver Springer, 1981.

[19] Leemis, L. M. *Reliability: Probability Models and Statistical Methods*, Prentice Hall , Englewood Cliffs, NJ, USA, 1995.

[20].Peter.J.Smith.Analysis of Failure and Survival data.2002.Champan and Hall/CRC