Information Theoretic Measures Based On Record Statistics

A thesis submitted to

DELHI TECHNOLOGICAL UNIVERSITY

for the award of the degree of

DOCTOR OF PHILOSOPHY

in

MATHEMATICS

by

RITU GOEL

under the supervision of

Prof. H.C.Taneja

and

Dr. Vikas Kumar

DEPARTMENT OF APPLIED MATHEMATICS

DELHI TECHNOLOGICAL UNIVERSITY

(FORMERLY DELHI COLLEGE OF ENGINEERING)

DELHI-110042, INDIA.

April, 2019 **Enroll. No. : 2K14/PhD/AM/03**

DECLARATION

I, **Ritu Goel** , student of Ph.D hereby declare that the research work reported in this thesis entitled “**Information Theoretic Measures Based On Record Statistics**" submitted by me to the Delhi Technological University, Delhi for the award of the degree of *Doctor of Philosophy in Mathematics* has been carried out by me under the supervision and guidance of Dr. H.C.Taneja, Professor in Department of Applied Mathematics, Delhi Technological University, Delhi and Dr. Vikas Kumar, Assistant Professor in UIET, Maharshi Dayanand University, Rohtak.

The research work embodied in this thesis is my original research and to the best of my knowledge has not been submitted earlier in part or full or in any other form to any other University or Institute for the award of any degree or diploma.

(Ritu Goel)
Department of Applied Mathematics
Delhi Technological University
Delhi-110042.

CERTIFICATE

This is to certify that the thesis entitled **"Information Theoretic Measures Based On Record Statistics"** submitted by **Ritu Goel** in the Department of Applied Mathematics, Delhi Technological University, Delhi, India for the award of degree of *Doctor of Philosophy in Mathematics*, is an original contribution with existing knowledge and faithful record of research work carried out by her under our guidance and supervision. We have gone through the work reported in this thesis and that, in our opinion, it is fully adequate in scope and quality as a thesis for the degree of Doctor of Philosophy.

To the best of our knowledge the work reported in this thesis is original and has not been submitted partially or fully to any other university or institution in any form for the award of any degree or diploma.

(Prof. H.C.Taneja)
Supervisor
Department of Applied Mathematics
Delhi Technological University
Delhi.

(Dr. Vikas Kumar)
Co-Supervisor
University Instt.of Engg. & Tech.
Maharshi Dayanand University
Rohtak.

(Dr. Sangita Kansal)
Professor & Head
Department of Applied Mathematics
Delhi Technological University
Delhi.

ACKNOWLEDGEMENTS

I have received encouragement and support from lot of people during my Ph.D. work. I would never have been able to complete this without their guidance and support and I think it is the time to express my gratitude to all of them.
First and foremost I would like to thank the almighty for opportunities and blessings bestowed upon me to reach this place.

I wish to express my sincere gratitude to my supervisors Prof. H.C.Taneja, Department of Applied Mathematics, Delhi Technological University, Delhi and Dr. Vikas Kumar, University Institute of Engineering and Technology, Maharshi Dayanand University, Rohtak, for their inspiring guidance and support during my research work, and for giving me the freedom to explore my own ideas. It is indeed a great pleasure for me to work under their supervision. I appreciate all their contributions of time and ideas to make my Ph.D. experience productive and stimulating. The joy and enthusiasm they have for their research was contagious and motivational for me.

I sincerely thank to Dr. Sangita Kansal, Professor and Head, Department of Applied Mathematics, DTU, for providing me the necessary facilities and valuable suggestions during the progress of the work.

I extend my sincere thanks to Prof. H. C. Taneja as a Dean (Academic, PG), DTU for his everlasting support and guidance. My sincere thanks to all faculty members of the Department of Applied Mathematics and other Departments of DTU for their constant support and encouragement.

I gratefully acknowledge the academic branch and administration of DTU for providing the environment and facilities to carry out my research work. Thanks are also due to office staff of Department of Applied Mathematics for their all kind of support.

I would like to thank Dr. Richa Thapliyal and Mr.Luckshay Batra for valuable discussions on the subject matter and for their support. I express my thanks to all research scholars of the Department of Applied Mathematics.

My time at DTU was made enjoyable due to my friends and groups that became a part of my life. I am thankful for time spent with friends Payal, Ajay Kumar and Mridula Mundalia.

I wish to record my profound gratitude to my parents who have provided me all kinds of support and help for my academic achievements, and for their constant love and care. I specially thank to my mother for taking care of every need of mine. I would like to express my thanks to my brothers, my sister and other family members for their heartiest cooperation and affection. I would like to extend my warm regards to my father-in-law and mother-in-law for their never ending support and love. I would like to thank my husband Dr. Amit Jindal for making everything possible for me and without his support this work could not have been completed and special thanks to my little princesses Thiwishaa and Prisha for patiently giving me maximum time to complete this work.

I want to express my thanks to all those who have not been mentioned here but supported, encouraged and inspired me during my Ph.D. work.

(RITU GOEL)

Dedicated to

My Family

Contents

Preface

The various information theoretic measures for example Shannon entropy measure [104] and its various additive generalizations like Renyi entropy [99], and Varma entropy [122], and non-additive generalizations like Havrda and Charvat entropy measure [56] have applications in different fields like statistics, physics, electronics etc. All these are based on single random variable. So by taking idea from this, we try to explore the behaviour of various information theoretic measures when these are applied on the sequence of record random variables and on the sequence of k-record random variables. In this thesis we study generalized Varma entropy measure, Kerridge inaccuracy measure, Kullback-Leibler discrimination measure, cumulative residual inaccuracy measure and entropy measure for past lifetime for the sequence of record values. We introduce these measures for record values and study some characterization results based on them.

This thesis includes seven chapters including the first chapter which is on introduction and literature survey. The thesis is organized as follows:

In **Chapter 2**, we have considered a measure of past entropy based on Shannon [104] entropy measure for n^{th} upper k-record value. A characterization result for the measure under consideration has given. We have discussed some basic properties of the proposed measure. Also we have constructed some bounds to the proposed past entropy measure for n^{th} k-records. The work reported in this chapter has been published in the paper entitled, **Measure of Entropy for Past Lifetime and k-Record Statistics** in *Physica A*, 2018, 503, 623-631.

In **Chapter 3**, we have introduced a measure of inaccuracy between distributions of the n^{th} record value and parent random variable and discussed some properties of it. It has also been shown that the proposed inaccuracy measure characterizes the distribution of parent random variable uniquely. Measure of inaccuracy for some specific distributions has also been studied. F^{α} distributions are equally important, so keeping this in mind we have also studied inaccuracy measure for F^{α} distributions. The part of the work reported in this chapter has been published in the paper entitled, **Kerridge Measure of Inaccuracy for Record Statistics** in *Journal of Information and Optimization Sciences*, 2018, 39(5), 1149-1161 and some work has been presented in the **International Conference on interdisciplinary Mathematics, Statistics and Computational Techniques** held at Manipal University, Jaipur, Dec 22-24, 2016.

In **Chapter 4**, we have studied a measure of inaccuracy between n^{th} upper k-record value and m^{th} upper k-record value. A simplified expression for the proposed inaccuracy measure has also been derived to find the inaccuracy measure for some specific probability distributions. We have also shown that the proposed inaccuracy measure characterizes the underlying distribution function uniquely. Further we have considered residual measure of inaccuracy between two k-record values and given a characterization result for that. The results reported in this chapter have been published in the paper entitled, **Measure of Inaccuracy and k-Record Statistics** in *Bulletin of Calcutta Mathematical Society*, 2018, 110 (2), 151-166 and some work has been presented in **National Seminar on Recent Developments in Mathematical Sciences** held at MDU, Rohtak, Mar 07-08, 2017.

In **Chapter 5**, we have provided an extension of cumulative residual inaccuracy as suggested by Taneja and Kumar [116] to k-record values. We have studied some properties of this measure. Also we have studied some stochastic ordering and have found the proposed measure for some of the distributions which occur often in many realistic situations and have applications in various fields of science and engineering. The work reported in this chapter is communicated under the title, **Cumulative Residual Inaccuracy Measure for k-Record Values** and some work has been presented in **International Conference on Recent Advances in Pure and Applied Mathematics** held at Delhi Technological University, Delhi, Oct 23-25, 2018.

In **Chapter 6**, we have provided an extension of Kullback-Leibler [66] information measure to k-record values. The distance between two k-record distributions of residual lifetime has been found. We have found the measure of discrepancy between n^{th} k-record value and m^{th} k-record value. Also keeping the record times fixed, we have derived the distance between k-record value and l-record value. We have also studied some properties of the measure proposed and a characterization result for that. The work reported in this chapter is communicated under the title, **A Measure of Discrimination Between Two Residual Lifetime Distributions For k-Record Values** and some work has been presented in **International Research Symposium on Engineering and Technology** held at Singapore, August 28-30, 2018.

In **Chapter 7**, we have considered and studied a generalized two parameters entropy based on Varma's entropy [122] function for k-record statistic. A general expression for this entropy measure of k-record value has been derived. Furthermore based on this, we have proposed a generalized residual entropy measure for k-record value and proved a characterization result that the proposed measure determines the distribution function uniquely. Also, an upper bound to the dynamic generalized entropy measure has been derived. The part of the work reported in this chapter has been communicated under the title, **On Generalized Information Measure of Order (α, β) and k-Record Statistics**.

In the last we have presented the conclusion of the work reported in this thesis and further scope of work.

Chapter 1

Introduction And Literature Survey

1.1 An Outline Of Information Theory

Information Theory is a branch of applied mathematics which deals with the problems like information processing, information storage, information retrieval, information utilization and decision-making. Basically it deals with all the theoretical problems which come across in the transmission of information over the communication channels. Information Theory has found applications in various fields like electrical engineering, financial mathematics, statistical modeling and image processing etc. Although the first attempt in this direction was made by Nyquist [90] and Hartley [55] considering the entropy measure for equally probable events, yet the theoretical foundation for all these developments dates backs to the work of Shannon [104] and others in the mid of the 20th Century which led to the development of information theory as a field of mathematics. The theory basically considers following three fundamental questions:

- **Compression:** How much can data be compressed (coded) so that another person can recover an identical copy of the uncompressed data?
- **Lossy data compression:** How much can data be compressed so that another person can recover an approximate copy of the uncompressed data?
- **Channel capacity:** How quickly reliable communication is possible from the source to destination through a noisy medium?

It is concerned with the mathematical laws governing systems designed to communicate or manipulate information. Being an electrical engineer Shannon's goal was to get maximum line capacity with minimum distortion. He was interested in the technical problems of high -fidelity transfer of message rather than semantic meaning of a message or its pragmatic effect on the listener.The second half of the 20th Century was characterized by the tremendous development of systems in which the transmitted information (analog signal) is coded in a digital form. By this coding the real nature of the information signal becomes secondary, that is, the same system can transmit simultaneously signals of very different nature: data, audio, video etc. This development has been made possible by use of more and more powerful integrated circuits.

In spite of the fact that Shannon presented entropy as a measure of uncertainty in the communication theory, the measure has kept on finding diverse applications in a variety of disciplines including mathematics, physics, biological sciences, pattern recognition etc, refer to [58, 101, 108, 119]. Another important area where information theoretic measures have found applications is that of reliability. Many researchers have studied the information-theoretic measures based on lifetime distributions of a system e.g. Ebrahimi [36], Ebrahimi and Kirmani [40], Asadi and Ebrahimi [11], Belzuence et al. [21], Nanda and Paul [88] and Kumar et al. [69, 70]. Another area of application of information theory which has drawn the attention of researchers is order statistics and record values. Many researchers have worked in the field connecting information theory and order statistics, refer to, Baratpour [19, 20], Ebrahimi et al. [42] and Asadi and Zarezadeh [127], Park [92], Ebrahimi et al. [42], Thapliyal and Taneja [120] and Thapliyal et al. [121] . Many have done work for the information measures contained in the sequence of record values. The information measures for record values have been investigated by several authors, including Zahedi and Shakil [126], Baratpour et al. [19] and Arghami and Abbasnejad [10]. To get rid of some of the difficulties in occurrence of record data, Dziubdziela and Kopocinski [35] introduced the model of k-record statistics. Instead of observing the sequence of largest values as in case of record values, he observed the sequence of k^{th} largest values. The field of k-record values has become the subject of interest for many researchers and mathematicians refer to Ahmadi and Doostparast [6], Ahmadi and Mohtashmi [7]. Motivating by this, in this thesis, we study the information theoretic measures for the record values and further for k-record values.

1.2 Entropy Measure And Its Generalizations

1.2.1 Shannon's Entropy

Shannon introduced a measure of information or entropy for a general finite complete probability distribution and gave a characterization theorem for the entropy measure introduced. Entropy as defined by Shannon [104] and added upon by other physicists is closely related to thermodynamical entropy. Infact, Shannon borrowed the idea of entropy from the second law of thermodynamics, according to this law the universe is winding down from an organized state to chaos, moving from predictability to uncertainty. Entropy is a measure of randomness. How much information a message contains is measured by the extent it combats entropy. The less predictable the message, the more information it carries. Consider the random variable $X = \{X_1, X_2, ...X_n\}$ with probability distribution $P = \{p_1, p_2, ..., p_n\}$, then the Shannon's entropy is defined as

$$H(P) = -\sum_{i=1}^{n} p_i \log p_i \ , \ \ 0 \le p_i \le 1, \ \ \sum_{i=1}^{n} p_i = 1 \tag{1.2.1}$$

Here it is assumed that $0 \log 0 = 0$; and normally the base of the logarithm is taken as 2, and then, the units are *'bits'* a short of the term *'binary digit'*. Shannon entropy provides the measure of average uncertainty associated with the outcome of the experiment or a measure of information conveyed through the knowledge of the probabilities associated with the events. It satisfies following important properties which are usually considered desirable for a measure of uncertainty defined in terms of probability distributions:

1. **Non-negativity:** $H(P)$ is always non-negative, that is,

$$H(P) = -\sum_{i=1}^{n} p_i \log p_i \ge 0 \ . \tag{1.2.2}$$

   The result is obvious since $-p_i \log p_i \ge 0$ for all i. It becomes zero, if one $p_i = 1$ and rest are zeros.

2. **Maxima:** $H(p_1, p_2, \ldots\ldots, p_n) \le \log n$, with equality when $p_i = \frac{1}{n}$ for all i.

3. **Continuity:** $H(p_1, p_2, \ldots\ldots, p_n)$ is a continuous function of p_i's, that is, a slight change in the probabilities p_i's results in the slight change in the uncertainty measure also.

4. **Symmetry:** $H(p_1, p_2, \ldots\ldots, p_n)$ is a symmetric function of p_i's, that is, it is invariant with respect to the order of the outcomes.

5. **Grouping (or, Branching) Property:**

$$H\{p_1, p_2, p_3, \cdots, p_n\} = H\{p_1 + \cdots + p_r,\ p_{r+1} + \cdots + p_n\} + (p_1 + \cdots + p_r) \times$$

$$H\left(\frac{p_1}{\sum_{i=1}^{r} p_i}, \cdots, \frac{p_r}{\sum_{i=1}^{r} p_i}\right) + (p_{r+1} + \cdots + p_n) H\left(\frac{p_{r+1}}{\sum_{i=r+1}^{n} p_i}, \cdots, \frac{p_n}{\sum_{i=r+1}^{n} p_i}\right)$$

for $r = 1, 2, \cdots n-1$.

6. **Additivity**: If $P = (p_1, p_2, \ldots\ldots, p_n)$ and $Q = (q_1, q_2, \ldots\ldots, q_n)$ are two independent probability distributions, then

$$H(P \bullet Q) = H(P) + H(Q),$$

where $P \bullet Q$ is the joint probability distribution, that is, for two independent distributions entropy of the joint distribution is the sum of the entropies of the two marginal distributions.

Corresponding to (1.2.1), for a continuous probability distribution $f(x)$, the measure of uncertainty is defined as

$$H(X) = -E[\log f(X)] = -\int_0^\infty f(x) \log f(x) dx\,. \tag{1.2.3}$$

In general, this measure is termed as **differential entropy**, for more details refer to McEliece [80]. Unlike the uncertainty measure (1.2.1) defined for discrete random variable which is always positive, the differential entropy can take negative value as well. The measure (1.2.3) has also been extensively studied by many researchers.

1.2.2 Generalizations

We have seen that Shannon entropy satisfies a number of useful properties like non-negativity, continuity, symmetry, additivity, grouping, etc. Many of the researchers have used some of these properties as axioms to characterize the Shannon entropy. The most intuitive and compact axioms for characterizing the Shannon entropy function have been given by Khinchin [65], which are known as the Shannon-Khinchin axioms. Many other researchers have also characterized Shannon entropy using different set of axioms. For some further results on characterization and the algebraic properties of Shannon entropy refer to Aczel and Darocozy [3].

Though Shannon's entropy is at the focus in information theory, yet the idea of information is so rich enabling no single definition that will have the capacity to measure information legitimately. Hence, many researchers presented the parametric group of entropies as a mathematical generalization of Shannon's entropy. These entropies are functions of some parameters and tend to Shannon entropy when these parameters approach their limiting values. For the first time, this was done by Renyi [99] who characterized a scalar parametric entropy as entropy of order α, which is additive in nature and includes Shannon entropy as a limiting case given as

$$H^{\alpha}(P) = \frac{1}{1-\alpha} \log \left\{ \sum_{i=1}^{n} {p_i}^{\alpha} \right\}; \ \alpha \neq 1, \ \alpha > 0. \qquad (1.2.4)$$

The additional parameter α makes it more sensitive to the shape of probability distributions. When $\alpha \to 1$, the Renyi entropy becomes Shannon entropy and it is substantially more versatile because of the parameter α, permitting several measurements of uncertainty with in a given distribution. Analogous to (1.2.4), the Renyi entropy for continuous random variable X, is given as

$$H^{\alpha}(X) = \frac{1}{1-\alpha} \log \left\{ \int_{0}^{\infty} f^{\alpha}(x) dx \right\}; \ \alpha \neq 1, \ \alpha > 0. \qquad (1.2.5)$$

Another important generalization of Shannon entropy is given by Varma [122] who introduced a entropy measure, which is again additive in nature as

$$H_\beta^\alpha(X)=\frac{1}{\beta-\alpha}\log\left\{\int_0^\infty f^{\alpha+\beta-1}(x)dx\right\},\quad \beta-1<\alpha<\beta,\beta\geq 1. \tag{1.2.6}$$

It reduces to Renyi entropy measure (1.2.5) when $\beta=1$ and Shannon entropy measure when $\beta=1,\ \alpha\to 1$.

In some realistic situations, the entropy measures which are additive in nature are not able to justify the behaviour observed. So keeping this in mind, various researchers introduced the non-additive entropy measures. In 1967, Havrda and Charvat [56] introduced the entropy measure, non-additive in nature as

$$H^\omega(P)=\frac{1}{1-\omega}\left[\sum_{i=1}^n p_i^\omega-1\right],\ \ \omega\neq 1,\ \omega>0. \tag{1.2.7}$$

Continuous analogous of (1.2.7), is given as

$$H^\omega(X)=\frac{1}{1-\omega}\left[\int_0^\infty f^\omega(x)dx-1\right],\ \ \omega\neq 1,\ \omega>0. \tag{1.2.8}$$

Also other than the entropy measures mentioned above there are several entropy measures available in the literature invented by various researchers refer to e.g. Kapur [60, 61], and Sharma and Taneja [105].

1.3 Relative Measure Of Discrimination And Measure Of Inaccuracy

Since the entropy measures have been considered as measure of information contained in a given probability distribution, it is normal to examine measures which allow one to evaluate the amount of information shared between two probability distributions or how close two distributions are from each other. One of the most important and useful measure of distance was given by Kullback and Leibler [66]. If $P=\{p_1,p_2,\ldots,p_n\}$ is the actual probability distribution associated with the outcomes $X=\{X_1,X_2,\ldots,X_n\}$ and $Q=\{q_1,q_2,\ldots,q_n\}$ is the predicted (or reference) distribution associated with the same experiment. Also $p_i\geq 0,\ q_i\geq 0$ and $\sum_{i=1}^n p_i=\sum_{i=1}^n q_i=1$, then Kullback's measure of relative information [66] is given by

$$H(P/Q) = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i} \,. \tag{1.3.1}$$

It is assumed that whenever $q_i = 0$, the corresponding p_i is also zero and $0 \log \frac{0}{0} = 0$. The different authors gave this measure many different names such as measure of discrimination, cross entropy, directed divergence, relative information etc. This measure plays an important role in information theory. It quantizes the discrimination between two probability distributions and we observe that $H(P/Q) = 0$, when $p_i = q_i \ \forall \, i$. Some generalizations of the measure have been studied by Taneja and Kumar [116], Taneja [113] and Kapur [60, 61].

Another important concept is that of inaccuracy measure. In making a statement about the probabilities of different outcomes in performing an experiment, generally there is possibility of two types of errors, to be specific, one due to the insufficient information and other from inaccurate information. Shannon entropy measure is capable to explain the error resulting from ambiguity only. The measure which considered the both type of errors was given by Kerridge [63] as

$$H(P;Q) = -\sum_{i=1}^{n} p_i \log q_i \,, \tag{1.3.2}$$

where q_i is the predicted probability and p_i is the actual probability of an outcome. This is a generalization of Shannon entropy [104] in the sense that when $q_i = p_i$ for all $i's$, then (1.3.2) reduces to (1.2.1), the Shannon entropy measure. The measure of information, discrimination and inaccuracy are connected as

$$H(P;Q) = H(P) + H(P/Q) \,, \tag{1.3.3}$$

that is, inaccuracy is the sum of entropy and discrimination. For a continuous random variable Nath [83] extended Kerridge's inaccuracy as

$$H(f;g) = -\int_0^\infty f(x) \log g(x) dx \,, \tag{1.3.4}$$

Also he discussed some properties of it. Here $f(x)$ denotes the actual distribution and $g(x)$ denotes the predicted distribution.

Corresponding to the measure (1.3.1), the measure for relative information for continuous random variable is given as

$$H(f/g) = \int_0^\infty f(x) \log \frac{f(x)}{g(x)} dx \,. \tag{1.3.5}$$

The Shannon's entropy, Kullback-Leibler's relative information and Kerridge's inaccuracy are the three classical measures of information associated with one and two probability distributions.

1.4 Some Basic Concepts In Reliability Theory

Consider a random variable X representing the lifetime of a system or a component and $F(x)$ as its distribution function. There are many other functions like survival function or reliability, Hazard rate function, reverse hazard rate function, mean residual life function, mean inactivity time, which describe the distribution of the random variable X completely. Next, we explain all these.

1.4.1 Reliability

Reliability or the survival function is defined as the probability that a given component or a system will perform its required function without failure for a given period of time when used under stated operating conditions. Consider a random variable X with pdf $f(x)$ and the distribution function $F(x)$ representing the lifetime of a component, then reliability is

$$\overline{F}(x) = \Pr(X > x) = \int_x^\infty f(u)du \,,$$

Here $\bar{F}(x) = 1 - F(x)$ is simply the *survival function*. It should be noted that $\bar{F}(x)$ is a decreasing function of x with $\bar{F}(0) = 1$ and $\lim_{x\to\infty} \bar{F}(x) = 0$. We can obtain the probability density function $f(x)$ of X from its survival function $\overline{F}(x)$ by the relationship

$$f(x) = -\frac{d}{dx}\overline{F}(x).$$

1.4.2 Hazard Rate Function

The *hazard rate function*, also known as the *conditional failure rate* , is a non-negative function given as

$$\lambda(x) = \lim_{\Delta x \to 0} \frac{P(x < X < x + \Delta x | X \geq x)}{\Delta x} .$$

Above expression represents the conditional probability of an item failing in the interval $(x, x+\Delta x)$ provided it has not failed by time x.

It is easy to see that for a continuous random variable X it is given by

$$\lambda(x) = \frac{f(x)}{\overline{F}(x)} = -\frac{d}{dx} \log \overline{F}(x). \tag{1.4.1}$$

In the discrete domain, with non-negative integral support, Xekalaki [125] defines the failure rate for a random variable X, as

$$\lambda(x) = \frac{P(X = x)}{P(X \geq x)} . \tag{1.4.2}$$

If X represents the lifetime of a component, then $\lambda(x)$ is the probability that the component will fail at $X = x$ given that it has not failed up to the time before x. The units of $\lambda(x)$ are probability of failure per unit of time, distance or cycle. In reliability analysis, a life distribution can be classified according to the shape of its hazard rate function $\lambda(x)$. Taking the bathtub curve, the early failure period has a decreasing hazard function as time goes by; the useful life period has a constant hazard function, and the wear-out period has an increasing hazard function. The hazard rate function and survival function $\overline{F}(.)$ holds the relationship

$$\overline{F}(x) = \exp\{-\int_0^x \lambda(t) dt\} . \tag{1.4.3}$$

1.4.3 Reverse Hazard Rate Function

The concept of reversed hazard rate is of great interest for many researchers, refer to Keilson and Sumita [64]. It is useful especially in survival analysis and reliability. The *reversed hazard rate* for a non-negative random variable X is given as

$$\mu(x) = \lim_{\Delta x \to 0} \frac{P(x - \Delta x < X < x | X \leq x)}{\Delta x} . \tag{1.4.4}$$

For a continuous random variable X which denotes the lifetime of a component, the reversed hazard rate is given as

$$\mu(x) = \frac{f(x)}{F(x)} = \frac{d}{dx} \log F(x) .$$

Here $f(x)$ and $F(x)$ are the probability density function and the distribution function of the random variable X. For the discrete random variable with non-negative integral support reversed hazard rate function is given as

$$\mu(x) = \frac{P(X = x)}{P(X < x)} . \tag{1.4.5}$$

We can also compute the distribution function from reversed hazard rate function uniquely using the relation

$$F(x) = \exp\left(- \int_x^\infty \mu(t)dt \right) .$$

We can see that the hazard rate and reversed hazard rate functions are related as

$$\mu(x) = \frac{\lambda(x)\overline{F}(x)}{F(x)} . \tag{1.4.6}$$

Finkelstein [45] has given that

$$\mu(x) = \frac{\lambda(x)}{\exp(-\int_0^x \lambda(t)dt) - 1} . \tag{1.4.7}$$

The reversed hazard rate function has found applications in forensic sciences to know the exact time of failure of a system or a unit. For more details, one can refer to Block et al. [23], Di Crescenzo [30], Gupta and Nanda [53], Gupta and Wu [54], Nair et al. [85] and Sengupta et al. [102].

1.4.4 Mean Residual Lifetime Function

While discussing reliability theory, the *mean residual life* (MRL) of a system or a component is another important aspect to discuss. It provides an idea of how long a device of any particular age can be expected to survive. It provides an idea to improve the average lifetime of a system.

For a continuous random variable X with $E(X) < \infty$, the *mean residual life function* is defined as

$$\delta(t) = E[X - t | X > t] = \frac{\int_t^\infty \overline{F}(x)dx}{\overline{F}(t)} \,. \tag{1.4.8}$$

If we have some idea about the expected duration for which the component under consideration will continue to work, then it becomes easy to replace or to re-schedule that component. While constructing the maintenance policies, this proves to be more useful than the failure rate. To study various properties and application of mean residual life function one can refer to Barlow and Proschan [18], Swartz [111] and Muth [82].

The relation between survival function and mean residual life function is given as

$$\overline{F}(t) = \frac{\delta_F(0)}{\delta_F(t)} \exp\left[-\int_0^t \frac{dx}{\delta_F(x)}\right]. \tag{1.4.9}$$

Further failure rate is connected with the mean residual life function as

$$\lambda_F(t) = \frac{\delta_F'(t) + 1}{\delta_F(t)} \,. \tag{1.4.10}$$

Several researchers have studied the characterizations problems of various probability models based on the mean residual life (MRL) function, refer to Sunoj et al. [110] and Sullo and Rutherford [108] .

1.5 Hazard Models

In this section we discuss two types of dependence structures between two probability distributions; one the proportional hazard model, and second, the proportional reversed hazard model which have been extensively used in survival analysis.

1.5.1 Proportional Hazard Model

The proportional hazard model was introduced by Cox [28] as a dependence structure among two distributions. In literature, this model has been used to imitate failure time data. This model is commonly known as Cox PH Model, but it was basically introduced by Lehmann [74]. It is equally applicable for both discrete and continuous random variables. It is quite useful for estimating the risk of failure associated with a vector of covariates.

Consider two non-negative continuous random variables X and Y with the same support denoting the time to failure of two systems with hazard rates $\lambda_F(x) = \frac{f(x)}{\bar{F}(x)}$ and $\lambda_G(x) = \frac{g(x)}{\bar{G}(x)}$ respectively, and if

$$\lambda_G(x) = \beta \lambda_F(x), \tag{1.5.1}$$

where β is a positive constant, then this model is called proportional hazard model (PHM).

We can easily see that (1.5.1) is equivalent to

$$\bar{G}(x) = [\bar{F}(x)]^{\beta},\ \beta > 0. \tag{1.5.2}$$

This model is quite useful from application point of view . Many researchers have found its applications in various fields such as reliability, medicine, survival analysis, economics etc. Many have worked on the problems of characterization of specific probability distributions using information theoretic measures under the consideration of proportional hazards model refer to Ebrahimi and Kirmani [40], Nair and Gupta [84] and Kumar [117].

Also there are some situations where the hazard rates of two random variables are not proportional uniformly over the whole time interval, but may be proportional differently in different intervals. To deal with such type of situations , Nanda and Das [86] introduced the *dynamic proportional hazard model* (DPHM) and studied their properties for different aging classes. The dynamic proportional hazard model can be obtained by proportional hazard model just by replacing β by some non-negative function of some parameter t, then the model is defined as

$$\lambda_G(x) = \beta(t)\ \lambda_F(x),\ \forall\, t > 0, \tag{1.5.3}$$

which considers different proportionality for different time intervals.

1.5.2 Proportional Reversed Hazard Model

Similar to proportional hazard model to analyze the failure time data Gupta et al. [52] introduced another model called the *proportional reversed hazard model* (PRHM). Later on Sengupta et al. [102] proved that in comparison to proportional hazard model, proportional reversed hazard model gives a better fit for some data set.

Consider two non-negative continuous random variables X *and* Y *with the same support and reversed hazard rates* $\mu_X(x) = \frac{f(x)}{F(x)}$ *and* $\mu_Y(x) = \frac{g(x)}{G(x)}$ *respectively, and if*

$$\mu_Y(x) = \beta \; \mu_X(x) \; , \;\; \beta > 0 \tag{1.5.4}$$

then the model is called proportional reversed hazard model (PRHM).
Alternatively, this model is similar to

$$G(x) = [F(x)]^{\beta} \; , \tag{1.5.5}$$

where $F(x)$ represents the baseline distribution function and $G(x)$ represents the reference distribution function.

When it comes to the analysis of right truncated data, proportional reversed hazard model is more useful than proportional hazard model . Various authors have studied properties and the comparison of both the models i.e. proportional hazard model and proportional reverse hazard model, refer to Gupta and Gupta [49] and Gupta and Wu [54]. Di Crescenzo [30] has studied some aging characteristics and stochastic orders properties of proportional reversed hazard model . Similar to dynamic proportional hazard model Nanda and Das [86] have proposed the *dynamic proportional reversed hazard model* (DPRHM) which is defined as

$$\mu_Y(x) = \beta(t) \; \mu_X(x) \;\; , \tag{1.5.6}$$

which considers different proportionality for different time intervals.

1.6 Dynamic Information Theoretic Measures

In various fields like reliability, survival analysis, actuary etc., the lifespan of a system or a component is of prime importance. In such a situation information theoretic measures as discussed earlier are not appropriate to measure the uncertainty of such systems. So in the present section we discuss two types of dynamic information theoretic measures :
(i) Residual information theoretic measures and
(ii) Past information theoretic measures
On the data which is left truncated, we apply the residual information theoretic measures and on the data which is right truncated, we apply past information theoretic measures.

1.6.1 Residual Information Theoretic Measures

If a system or a component has worked up to the time 't', then the remaining lifetime is called the residual lifetime of that system or component. Consider a random variable X which represents the lifetime distribution of a component, then the random variable $[X-t|X>t]$ denotes the *residual lifetime* of that component.

To determine the uncertainty about the remaining lifetime of a system which has already worked up to the time 't', the Shannon's differential entropy (1.2.3) is not appropriate. So to deal with such type of systems, Ebrahimi [36] introduced the concept of residual entropy. He proposed the dynamic measure of entropy based on Shannon entropy as

$$H(f;t) = -E[\log f_t(X_t)] = \int_t^{\infty} -f_t(x)\log f_t(x)dx, \tag{1.6.1}$$

where $f_t(x)$ denotes the probability density function of the random variable $X_t = [X-t|X>t]$ given as

$$f_t(x) = \begin{cases} \frac{f(x)}{\overline{F}(t)} & ; \quad \text{if } x>t \\ 0 & ; \quad \text{otherwise} \end{cases}$$

$H(f;t)$ basically provides the expected uncertainty about the predictability of remaining lifetime of the unit contained in the conditional density of $X-t$ given $X>t$.

Using the probability density function of residual lifetime in (1.6.1), we get

$$H(f;t) = \int_t^\infty -\frac{f(x)}{\bar{F}(t)}(x)\log\frac{f(x)}{\bar{F}(t)}dx. \tag{1.6.2}$$

It is easy to see that the residual entropy in terms of hazard rate function can be written as

$$\begin{aligned} H(f;t) &= -\int_t^\infty \frac{f(x)}{\bar{F}(t)}\log\frac{f(x)}{\bar{F}(t)}dx \qquad (1.6.3)\\ &= \log\bar{F}(t) - \int_t^\infty \frac{f(x)}{\bar{F}(t)}\log f(x)dx \\ &= 1 - \int_t^\infty \frac{f(x)}{\bar{F}(t)}\log\lambda_F(x)\ , \qquad (1.6.4) \end{aligned}$$

where $\lambda_F(x) = \frac{f(x)}{\bar{F}(x)}$ is the hazard rate function.

Ebrahimi [36] has also shown that the residual entropy measure (1.6.2) determines the underlying distribution function uniquely. After Ebrahimi [36] introduced the residual entropy measure, many researchers have studied various results concerning the Shannon residual entropy measure. Rajesh and Nair [96] considerd the discrete case and introduced the similar results. Asadi and Ebrahimi [11] have given characterization results using the relationship between dynamic entropy and mean residual life of a system. Belzunce et al. [21] have studied similar results for generalized residual entropy. Nanda and Paul [88] have studied some characterization results for distributions based on a generalized residual entropy function. For more details one can refer to Ebrahimi and Pellerey [41], Ebrahimi [37], Sankaran and Gupta [100] and Ebrahimi [38].

Asadi et al. [12], Abraham and Sankaran [2], Baig and Dar [15] and Abbasnejada et al. [1] have also obtained some other results in connection with the Renyi entropy and Varma entropy.

Next, analogous to Kullback-Leibler relative information measure (1.3.5), for two non-negative continuous random variables X and Y with the same supports that represent the lifetimes of two systems, Ebrahimi and Kirmani [39] proposed the measure of discrepancy between two residual-life distributions as

$$\begin{aligned} H(f/g;t) &= \int_t^\infty \frac{f(x)}{\bar{F}(t)} \log \frac{f(x)/\bar{F}(t)}{g(x)/\bar{G}(t)} dx \\ &= \log G(t) - H(f;t) - \int_t^\infty \frac{f(x)}{\bar{F}(t)} \log g(x) dx \,. \end{aligned} \tag{1.6.5}$$

$H(f/g;t)$ can be interpreted as a measure of distance between $f_t(x)$ and $g_t(x)$. Here $\bar{F}(t)$ represents the actual survival distribution and $\bar{G}(t)$ represents the reference distribution. This measure is quite useful in ordering and classification of survival function. Further results on this measure have been studied by Ebrahimi and Kirmani [40] and Asadi et al. [12]. Also we can see that for each fixed $t > 0,$ $H(f/g;t),$ has all the properties of the Kullback-Leibler discrimination information measure $H(f/g).$ Recently some extensions of (1.6.5) have given by Navarro et al. [89] and the same have been used to characterize some bivariate distributions. These distributions are also characterized in terms of proportional hazard rate models and weighted distributions.

1.6.2 Past Information Theoretic Measures

In many real life situations, it is not always necessary that uncertainty is related to the future but it can also be connected with the past. For example consider a system which is examined at some preassigned time intervals, and it is found dead at any time t, then the uncertainty of the system's life is related with the past. More specifically, consider a system which is examined at times $\omega, 2\omega, 3\omega,$ ... for some preassigned time ω, it is possible that at time $(n-1)\omega$ the system is functioning, but at time $n\omega$ the system is found to be down, where n is a positive integer. Then, if X represents the failure time of the system, then our interest lies in $[n\omega - X|X \leq n\omega]$ or in general ${}_tX = [t - X|X \leq t]$ which is known as the *inactivity time or the past lifetime*. For further details related to past lifetime, one can refer to Chandra and Roy [25, 26] and Kayid and Ahmad [62]. So keeping this in mind Di Crescenzo and Longobardi [31] have considered measure of past entropy over

$(0,\, t)$ given by

$$H^*(f;t) = -\int_0^t \frac{f(x)}{F(t)} \log \frac{f(x)}{F(t)} dx \qquad (1.6.6)$$

$$= 1 - \int_0^t \frac{f(x)}{F(t)} \log \mu_F(x) dx,$$

where $\mu_F(x)$ is the reversed hazard rate function of X. Analogous to $f_t(x)$, in case of residual lifetime, the probability density function of the past lifetime random variable $_tX$, is given by $\frac{f(x)}{F(t)} = f_t^*(x)$ for $X \leq t$. Some applications of past uncertainty measure is found in forensic sciences where the knowledge of exact time of failure is important, this type of measures are of added value. Using relationships among past entropy, reversed hazard rate and mean inactivity time, Kundu et al. [73] have given some characterization results. Nanda and Paul [87] have studied few ordering properties based on this measure. Further extending these results, Di Crescenzo and Longobardi [32] have given a measure of divergence which gives a distance between two past lifetimes distributions. The discrimination measure between past lifetimes is

$$H^*(f/g;t) = \int_0^t \frac{f(x)}{F(t)} \log \frac{f(x)/F(t)}{g(x)/G(t)} dx. \qquad (1.6.7)$$

Here $F(.)$ represents the true distribution and $G(.)$ represents the reference distribution function. Basically $H^*(f/g;t)$ can also be considered as a measure of distance between $G_t^*(x)$ and the true distribution $F_t^*(x)$. Further if X and Y satisfy the proportional reversed hazard model (PRHM) then $H^*(f/g;t)$ is constant and vice versa, refer to Di Crescenzo and Longobardi [32]. Further results on this in context to generalized measure of discrimination have been studied by Hooda and Saxena [57]. Also a measure of inaccuracy between two past lifetime distributions has been studied by Kumar et al. [68] which is given as

$$H^*(f,g;t) = -\int_0^t \frac{f(x)}{F(t)} \log \frac{g(x)}{G(t)} dx. \qquad (1.6.8)$$

1.7 Distribution Functions Based Information Theoretic Measures

Though entropy measure (1.2.1) by Shannon is quite important in information theory and various other fields. However for a continuous random variable, Shannon differential entropy (1.2.3) raises the following issue:

- It is density function based information theoretic measure. The density function of a random variable in general may or may not exist. This is the case when the cumulative distribution function (CDF) is not derivable, in that case it is not possible to define the differential entropy.

- The differential entropy can take a negative value in case of continuous random variable while Shannon entropy of a discrete random variable is always non-negative.

- Shannon entropy does not converge asymptotically to differential entropy when evaluated from samples of a random variable. For further details refer to Rao [94].

Rao [95] took in to account these issues and developed another measure of information or randomness as

$$\xi(X) = -\int_0^{\infty} \bar{F}(x)\log\bar{F}(x)dx\,, \tag{1.7.1}$$

where $\bar{F}(x) = 1 - F(x)$ is the survival function of a random variable X. This is called *Cumulative Residual Entropy (CRE)*. It is based on distribution function unlike Shannon entropy measure which is based on density function of the random variable X. It has following merits over the differential entropy:

- Cumulative residual entropy is always non-negative.
- It has a steady definition for both discrete and continuous domains.
- Cumulative residual entropy can be easily computed from sample data and these computations asymptotically converge to the true value.

Basically the objective was to introduce a new measure of randomness which makes use of distribution function rather than density function because the density function is the derivative of distribution function and hence it is more regular than the density function. Also in real life situations the distribution function is of more interest. For example, if we consider a random variable denoting the life span of a machine, then the event of interest is not whether the life span equals a specific instant, but rather whether the life span exceeds that instant. The definition by Rao [94] also protects the principle that the logarithm of the probability of an event should represent the information content in the event. Many properties of this measure have been studied by Rao [95]. Further he studied its applications in reliability engineering refer to Rao [95].
Asadi and Zohrevand [13] have proposed the cumulative residual entropy (CRE) for the residual lifetime distribution of a system as

$$\xi(X;t) = -\int_t^\infty \frac{\bar{F}(x)}{\bar{F}(t)} \log \frac{\bar{F}(x)}{\bar{F}(t)} dx\,. \tag{1.7.2}$$

Also, Di Crescenzo and Longobardi [34] proposed the uncertainty measure based on the failure distribution as

$$\xi(X) = -\int_0^\infty F(x) \log F(x) dx\,. \tag{1.7.3}$$

1.8 Order Statistics

The term "order statistics" was introduced by Wilks in 1942. However, the subject is much older. Early in the nineteenth century measures under consideration included the median, symmetrically trimmed means, the mid range, and related functions of order statistics. In 1818, Laplace obtained (essentially) the distribution of the r^{th}-order statistic in random samples and also derived a condition on the parent density under which the median is asymptotically more efficient than the mean. Other topics considered are of more recent origin: extreme-value theory and the estimation of location and scale parameters by order statistics. Suppose that $X_1, X_2, \cdots, X_n$ are independent and identically distributed observations from a distribution F_X, where F_X is differentiable with a density function f_X, which is positive in an interval and zero elsewhere. Order statistics of the sample is defined by the arrangement of $X_1, X_2, \cdots, X_n$ from the smallest to the largest denoted as

$X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$. It is well known that the p.d.f of i^{th} order statistics is

$$f_{i:n}(y) = \frac{1}{B(i, n-i+1)} (F_X(y))^{i-1} (1 - F_X(y))^{n-i} f_X(y),$$

where $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx,\ a > 0,\ b > 0$ is the beta Integral with parameters a and b (for details refer to Arnold et al. [9]).
The subject of order statistics deals with the properties and applications of these ordered random variables and of functions involving them. These statistics have been used in a wide range of problems like detection of outliers, characterizations of probability distributions, quality control and strength of materials (refer to David and Nagaraja [29]). In reliability theory, order statistics are used for statistical modelling. The m^{th} order statistics in a sample of size n represents the life length of a $(n-m+1)$-out-of-n system. For $m = 1$ and n, that is for sample minima and maxima, $(n-m+1)$-out-of-n system corresponds to series and parallel system respectively.
In 1990, Wong and Chen [124] calculated entropy of an ordered sequence. Park [92] obtained some recurrence relations for this entropy of order statistics. Ebrahimi et al. [41] explored Shannon entropy and some of its properties for order statistics. Arghami and Abbasnejad [10] explored information theoretic properties of Renyi entropy based on order statistics. Thapliyal and Taneja [120] studied dynamic residual Renyi entropy and dynamic cumulative residual entropy based on order statistics.

1.9 Record Values

The concept of record values was first introduced by Chandler [24]. With the concept of record values he also introduced the terms record times and inter record times. He also studied some of the interesting properties of record values and proved that for any random variable with any distribution, the expected value of the inter record time is infinite. Feller [44] also studied record values with respect to gambling problems. Record values sequences are of two types.
(i) Upper record values
(ii) Lower record values
Let $\{X_i,\ i \geq 1\}$ be a sequence of independent and identically distributed continuous random variables with distribution function $F(x)$ and probability density func-

tion $f(x)$. An observation X_j is called an *upper record value* if its value exceeds that of all previous observations. Thus X_j is an upper record if $X_j > X_i$ for every $i < j$. *A lower record value can be defined similarly.* One can get lower record values sequence from upper record value sequence by replacing $\{X_j\}$ by $\{-X_j\}$. Records are closely connected with the occurrence times of a corresponding non-homogeneous Poisson process.

Record Times: The times at which upper record values or lower record values appear are given by the random variables T_j which are called record times. $T_1 = 1$ with probability 1 and for $j \geq 2$, T_j = min $\{i : X_i > X_{T_{j-1}}\}$. The sequence of upper record values can thus be defined by $\{X_{T_j}\}$, j = 1, 2, 3 .....

Inter- Record Times: Let $\Delta_j = T_{j+1} - T_j$ and

$$\Delta_{(j)} = \acute{T}_{j+1} - \acute{T}_j, \;\; j = 1, 2, ...$$

then Δ_j and $\Delta_{(j)}$ are the upper and lower inter record times respectively.

To make more insight of this concept of record values following examples are helpful.

1. Suppose following are the ten observations from a given experiment:
$10, 12, 6, 15, 20, 18, 17, 5, 22, 3.$
The lower record values are: $10, 6, 5, 3.$
The upper record values are: $10, 12, 15, 20, 22.$

2. Consider a sequence of objects that may stop working when we apply some kind of stress on them and we want to find that minimum amount of stress under which these objects stop working. So we test the objects for minimal failure stress sequentially. We apply the stress on the first object until it gets fail and record this minimal stress as X_1. Now we test the next object for minimal failure stress and record it only if it is less than X_1 otherwise we consider the next object. In general, we will record stress X_n only if it is less than all the previous minimal stresses i.e. $X_n < min(X_1, X_2, ..., X_{n-1}), n > 1$. In this way we get a sequence of lower record values.

3. Consider the weighing of some objects on a scale missing its spring. When we place an object on the scale, needle indicates the correct weight but as its spring is missing, it can not return to zero back when the object is removed. If we weigh the various objects, only the weight greater than the previous ones can be recorded. In this way we get the sequence of upper record values.

The probability density function of upper record values (refer to Ahsanullah [8]) can be given as

$$f_j(x) = \frac{R^{j-1}(x)}{\Gamma(j)} f(x), -\infty < x < \infty. \tag{1.9.1}$$

where $r(x) = \frac{d}{dx}R(x) = \frac{f(x)}{1-F(x)}$, and $R(x) = -\ln(1-F(x))$, the function r(x) is the hazard rate.

Similarly p.d.f of j^{th} lower record value is given by

$$f_j(x) = \frac{H^{j-1}(x)}{\Gamma(j)} f(x), -\infty < x < \infty, \tag{1.9.2}$$

where $H(x) = -\ln F(x)$.

1.10 k-Record Values

The record model becomes inadequate in several situations like when the expected waiting time between two record values is very large. In those situations, k-record values are of great importance, see Kamps [59]. So, the concept of k-record values has been studied in the literature widely, refer to Berred [22], Fashandi and Ahmadi [43].

The model of k-record values was first introduced by Dziubdziela and Kopocinski [35]. They defined the k-record values in terms of k^{th} largest X yet seen, where k is any positive integer. Define $T_{1,k} = k$ and for $n \geq 2$,
$T_{n,k} = min\left\{ j : j > T_{n-1,k}, X_{j-k+1:j} > X_{T_{n-1,k}-k+1:T_{n-1,k}} \right\}$, $Y_{n,k} = X_{T_{n,k}-k+1:T_{n,k}}, n \geq 1$,
where $X_{i:n}$ represents the i^{th} order statistic in a sample of size n and $\{Y_{n,k,} \; n \geq 1\}$

represents the sequence of upper k-record values. The lower k-record values can be defined in a similar way. The probability density function of n^{th} upper k-record and the n^{th} lower k-record are given by

$$f_{n,k}(x) = \frac{k^n}{\Gamma(n)}[-\log \bar{F}(x)]^{n-1}[\bar{F}(x)]^{k-1}f(x) \tag{1.10.1}$$

and

$$f^L_{n,k}(x) = \frac{k^n}{\Gamma(n)}[-\log F(x)]^{n-1}[F(x)]^{k-1}f(x), \tag{1.10.2}$$

respectively, where $\Gamma(.)$ is the complete gamma function, refer to Arnold et al. [9]. The survival function of n^{th} upper k-record value and n^{th} lower k-record value is given as

$$\bar{F}_{n,k}(x) = \sum_{j=0}^{n-1} \frac{1}{j!} \left(\bar{F}(x)\right)^k \left(-k \ln \bar{F}(x)\right)^j \tag{1.10.3}$$

and

$$\bar{F}^L_{n,k}(x) = \sum_{j=0}^{n-1} \frac{1}{j!} \left(F(x)\right)^k \left(-k \ln F(x)\right)^j. \tag{1.10.4}$$

Ordinary record values can be obtained from k-records by putting $k = 1$.
To elaborate the concept of k-record statistics, consider the following data giving the average temperature of a city during the month of July measured for twenty years, refer to Arnold et al. [9]:
19.0, 20.1, 18.4, 17.4, 19.7, 21.0, 21.4, 19.2, 19.9, 20.4, 20.9, 17.2, 20.2, 17.8, 18.1, 15.6, 19.4, 21.7, 16.2, 16.4
For $k = 3$, the upper and lower 3-record respectively denoted by $R^U_{n,3}$, $R^L_{n,3}$ and 3-record times respectively denoted by $T^U_{n,3}$ and $T^L_{n,3}$ are :

n	1	2	3	4	5	6	7
$T^U_{n,3}$	3	5	6	7	10	11	18
$R^U_{n,3}$	18.4	19.0	19.7	20.1	20.4	20.9	21.0
$T^L_{n,3}$	3	4	12	14	16	19	20
$R^L_{n,3}$	20.1	19.0	18.4	17.8	17.4	17.2	16.4

The n^{th} k-record value represents the life length of a $k-$*out of* $-T_{n,k}$ system. The concept of record values and k-record values have been used in a wide range of problems like estimation of parameters and prediction of future records, characterizations of probability distributions, refer to Balakrishnan and Chan [16], Sultan et al. [107], Balakrishnan and Stepanov [17] and Su et al. [106].

1.11 Record Values And Information Theory

Apart from applications of record values and k-record values in different areas cited above, a huge literature is devoted to information theoretic measures based on record values and k-record values. Entropy properties of record values have been studied by several authors, refer to Zahedi and Shakil [126], Baratpour et al. [19, 20], Madadi and Tata [76], Razmkhan et al. [97], Asha and Chacko [14] and Kumar [71].

1.12 Motivation And Plan Of Work

Considering the importance and various applications of record values and k-record values in different fields, and in view of the above discussion and literature review, we were motivated to study information theoretic measures based on record values and k-record values. As the generalized entropy measures are useful due to flexibility provided by addition of parameters, so we have studied two parametric generalizations of Shannon entropy for k-record values and have also studied its dynamic version. Also, we have found considerable interest in studying inaccuracy measures between two record values and further between two k-record values and also a discrepancy measure between two k-record values, since these aspects were yet to be explored in case of record values. Also we have suggested the past entropy measure for k-record values, as it is a useful and interesting concept not explored so far. This thesis includes seven chapters including the current chapter on introduction and literature survey. The organization of the work reported is as follows:

In **Chapter 2**, we have introduced a measure of entropy for past lifetime based on Shannon's entropy measure [104] for n^{th} upper k-record value. A characterization result for the measure under consideration has been given. We have discussed some basic properties of the proposed measure. Also we have constructed some bounds to the proposed past entropy measure for n^{th} k-records. The work reported in this chapter has been published in the paper entitled, **Measure of Entropy for Past Lifetime and k-Record Statistics** in *Physica A*, 2018, 503, 623-631.

In **Chapter 3**, we have introduced a measure of inaccuracy between distributions of the n^{th} record value and parent random variable and have studied a characterization result. Measures of inaccuracy for some specific distributions have also been studied. The F^{α} or the power distributions are equally important, so keeping this in mind we have also studied inaccuracy measure for power distribution. The part of the work reported in this chapter has been published in the paper entitled, **Kerridge Measure of Inaccuracy for Record Statistics**, *Journal of Information and Optimization Sciences*, 2018, 39(5), 1149-1161 and some work has been presented in the **International Conference on interdisciplinary Mathematics, Statistics and Computational Techniques** held at Manipal University, Jaipur, Dec 22-24, 2016.

In **Chapter 4**, we have studied a measure of inaccuracy between n^{th} upper k-record value and m^{th} upper k-record value. A simplified expression for the proposed inaccuracy measure has also been derived to find the inaccuracy measure for some specific probability distributions. We have also shown that the proposed inaccuracy measure characterizes the underlying distribution function uniquely. Further we have considered residual measure of inaccuracy between distribution of n^{th} upper k-record values and parent distribution and have given a characterization result for that. The results reported in this chapter have been published in the paper entitled, **Measure of Inaccuracy and k-Record Statistics**, *Bulletin of Calcutta Mathematical Society*, 2018, 110 (2), 151-166 and some work has been presented in **National Seminar on Recent Developments in Mathematical Sciences** held at MDU, Rohtak, Mar 07-08, 2017.

In **Chapter 5**, taking in to account the importance of cumulative residual entropy (CRE) measures, we have provided an extension of cumulative residual inaccuracy, refer to Taneja and Kumar [116], to k-record values. We have studied some properties of this measure. Also we have discussed some stochastic ordering and have found the proposed measure for some of the distributions which occur often in many realistic situations and have applications in various fields of science and engineering. The work reported in this chapter is communicated under the title, **Cumulative Residual Inaccuracy Measure for k-Record Values** and some work has been presented in **International Conference on Recent Advances in Pure and Applied Mathematics** held at Delhi Technological University, Delhi, Oct 23-25, 2018.

In **Chapter 6**, we have provided an extension of Kullback Leibler [66] information measure to k-record values. The distance between two k-record distributions of residual lifetime has been found. We have found the measure of distance or discrepancy between n^{th} k-record value and m^{th} k-record value. Also keeping the record times fixed, we have derived the distance between k-record value and l-record value. We have also studied some properties of the measure proposed and a characterization result for that. The work reported in this chapter is communicated under the title, **A Measure of Discrimination Between Two Residual Lifetime Distributions For k-Record Values** and some work has been presented in **International Research Symposium on Engineering and Technology** held at Singapore, August 28-30, 2018.

In **Chapter 7**, we have considered and studied a generalized two parameters entropy for k-record statistic based on Varma's entropy [122] function. A simplified expression for this entropy measure has also been derived. Further based on this, we have proposed a generalized residual entropy measure for k-record value and have proved a characterization result. Also, an upper bound to the dynamic generalized entropy measure has been derived. The part of the work reported in this chapter has been communicated under the title, **On Generalized Information Measure of Order (α,β) and k-Record Statistics**.

In the last we have presented the conclusion of the work reported in this thesis and further scope of work, followed by bibliography and the list of publications.

Chapter 2

Measure Of Entropy For Past Lifetime And k-Record Values

2.1 Introduction

The measure of Shannon entropy which gives the average uncertainty contained in the probability density function $f(x)$ associated with the random variable X is not suitable as a measure of uncertainty for the remaining lifetime of a system which has already survived up to time t. This time period $'t'$ has great importance in economics, reliability, business and survival analysis. For such type of systems Ebrahimi [36] proposed the measure of uncertainty of the remaining lifetime $X_t = (X-t|X>t)$ as

$$H(f;t) = -\int_t^\infty \frac{f(x)}{\bar{F}(t)} \log \frac{f(x)}{\bar{F}(t)} dt, \;\; t>0. \tag{2.1.1}$$

In many realistic situations, uncertainty is not related to future only, this may refer to past also. Suppose some one has under gone a medical test at time t, to check for a certain disease and that test is found positive. Let X denotes the age when the patient was infected, then obviously $X<t$. Now the question is, how much time has elapsed since the patient had been infected by this disease. Such random time can be called *inactivity time or past lifetime* ${}_tX = [t-X|X<t]$, for fixed $t>0$.

The work reported in the present chapter has been published in the paper **Measure of Entropy for Past Lifetime and k-Record Statistics** in *Physica A*, 2018, 503, 623-631.

This gives the time elapsed from the failure of a component given that its lifetime is less than or equal to t. It is also called the reversed residual lifetime. Based on this idea Dicrescenzo and Longobardi [31] defined the measure of uncertainty for past lifetime distribution over $(0,t)$ as

$$\bar{H}(f;t) = -\int_0^t \frac{f(x)}{F(t)} \log \frac{f(x)}{F(t)} dx. \tag{2.1.2}$$

The measure (2.1.2) can be considered as dual to the measure (2.1.1) proposed by Ebrahimi [36]. The above measure has applications in forensic science. The measures of uncertainty for inactivity time, called past entropy, have been studied by various researchers, refer to Di Crescenzo and Longobardi [32,33] and Kumar et al. [68] and Kundu et al. [73].

Also record values originates in many realistic situations. Records are closely connected with the occurrence times of a corresponding non-homogeneous Poisson process and reliability theory. But this record model becomes inadequate in several situations when the expected waiting time between two record values is very large. In those situations, second or third highest values are of great importance, see Kamps [59]. So, the concept of k-record values has been studied in the literature widely, refer to Berred [22], Fashandi and Ahmadi [43].
These statistics have been used in a wide range of problems like estimation of parameters and prediction of future records, characterizations of probability distributions, refer to Balakrishnan and Chan [16], Sultan et al. [107], Balakrishnan and Stepanov [17] and Su et al. [106]. Entropy properties of record values have been investigated by several authors, refer to Zahedi and Shakil [126], Madadi and Tata [76], Razmkhan et al. [97] and Asha and Chacko [14].
In the present chapter we consider a measure of past entropy in the context of k-record values. The chapter is organized as follows: In Section 2.2, we propose the measure of past entropy and give a characterization result for the proposed measure in Section 2.3. In Section 2.4, we derive some bounds to the proposed measure and study its important properties like effect of linear transformation and stochastic ordering. In Section 2.5, we find the entropy measure for past lifetime for some specific distributions. Section 2.6 is devoted to conclusion.

2.2 Past Entropy For Upper k-Record

Corresponding to measure (2.1.2) proposed by Dicrescenzo and Longobardi [31], we define the measure of uncertainty for past life distribution over the interval $(0,t)$ for the distribution of n^{th} upper k-record value as follows:

$$\begin{aligned} \bar{H}(f_{n,k};t) &= -\int_0^t \frac{f_{n,k}(x)}{F_{n,k}(t)} \ln\left(\frac{f_{n,k}(x)}{F_{n,k}(t)}\right) dx \\ &= \ln F_{n,k}(t) - \frac{1}{F_{n,k}(t)} \int_0^t f_{n,k}(x) \ln f_{n,k}(x) dx \\ &= 1 - \frac{1}{F_{n,k}(t)} \int_0^t f_{n,k}(x) \ln r_{n,k}(x) dx, \end{aligned} \tag{2.2.1}$$

where $f_{n,k}(x)$ is the pdf of n^{th} upper k-record value given by (1.10.1) and $r_{n,k}(x) = \frac{f_{n,k}(x)}{F_{n,k}(x)}$ is the reversed hazard rate of $X_{n,k}$.
Differentiating (2.2.1) with respect to t both sides, we get

$$\begin{aligned} \frac{d}{dt}(\bar{H}(f_{n,k};t)) &= -r_{n,k}(t)\ln r_{n,k}(t) + \frac{f_{n,k}(t)}{F^2_{n,k}(t)} \int_0^t f_{n,k}(x) \ln r_{n,k}(x) dx \\ &= -r_{n,k}(t)\ln r_{n,k}(t) + r_{n,k}(t) \int_0^t \frac{f_{n,k}(x)}{F_{n,k}(t)} \ln r_{n,k}(x) dx. \end{aligned}$$

Using (2.2.1), we get

$$\frac{d}{dt}(\bar{H}(f_{n,k};t)) = r_{n,k}(t)\left(1 - \bar{H}(f_{n,k};t) - \ln r_{n,k}(t)\right). \tag{2.2.2}$$

Remark 2.2.1. *When we put $k = 1$, (2.2.1) defines the past entropy measure for usual records as*

$$\begin{aligned} \bar{H}(f_n;t) &= -\int_0^t \frac{f_n(x)}{F_n(t)} \ln\left(\frac{f_n(x)}{F_n(t)}\right) dx \\ &= 1 - \frac{1}{F_n(t)} \int_0^t f_n(x) \ln r_n(x) dx, \end{aligned}$$

where $f_n(x)$ is the pdf of n^{th} upper record value given by

$$f_n(x) = \frac{\left(-\ln \bar{F}(x)\right)^{n-1} f(x)}{\Gamma n}$$

and $r_n(x) = \frac{f_n(x)}{F_n(x)}$ is the reversed hazard rate function of X_n, the n^{th} record value.
Further for $n = k = 1$, the equation (2.2.1) gives the past entropy measure for parent

random variable. Also we can express Shannon differential entropy of n^{th} upper k-record value, in terms of the past entropy of n^{th} upper k-record value (2.2.1) and residual entropy of $X_{n,k}$

$$H(f_{n,k};t) = -\int_t^{\infty} \frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)} \ln\left(\frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)}\right) dx, \tag{2.2.3}$$

given as

$$H(f_{n,k}) = F_{n,k}(t)\bar{H}(f_{n,k};t) + \bar{F}_{n,k}(t)H(f_{n,k};t) + \beta(F_{n,k}(t), \bar{F}_{n,k}(t)), \tag{2.2.4}$$

where $\beta(p, 1-p) = -p\ln p - (1-p)\ln(1-p)$ denotes the Shannon entropy measure for a Bernoulli random variable. This follows as

$$\begin{aligned}
H(f_{n,k}) &= -\int_0^{\infty} f_{n,k}(x)\ln f_{n,k}(x)dx \\
&= -\int_0^{t} f_{n,k}(x)\ln f_{n,k}(x)dx - \int_t^{\infty} f_{n,k}(x)\ln f_{n,k}(x)dx \\
&= -F_{n,k}(t)\int_0^{t} \frac{f_{n,k}(x)}{F_{n,k}(t)}\ln f_{n,k}(x)dx - \bar{F}_{n,k}(t)\int_t^{\infty} \frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)}\ln f_{n,k}(x)dx \\
&= F_{n,k}(t)\bar{H}(f_{n,k};t) + \bar{F}_{n,k}(t)H(f_{n,k};t) - \bar{F}_{n,k}(t)\ln\bar{F}_{n,k}(t) - F_{n,k}(t)\ln F_{n,k}(t) \\
&= F_{n,k}(t)\bar{H}(f_{n,k};t) + \bar{F}_{n,k}(t)H(f_{n,k};t) + B(F_{n,k}(t), \bar{F}_{n,k}(t)).
\end{aligned}$$

The above result shows that the uncertainty contained in the failure time of a component or an item can be decomposed in to three parts :

- The uncertainty of whether the item has failed before or after time 't'.
- The uncertainty about the failure time in $(0,t)$ given that the item has failed before 't'.
- The uncertainty about the lifetime in (t,∞) given that the item has failed after 't'.

2.3 Characterization Result

Shannon information measure can be equal for two different distributions, so a distribution function can not be described by its Shannon entropy. A natural question arises that whether the proposed past measure of entropy determines the lifetime distribution $F(.)$ uniquely. In this section we study the condition under which the proposed measure determines the parent distribution uniquely. For that we use the lemma and theorem due to Gupta and Kirmani [51] which is stated as below.

Theorem 2.3.1. *Consider a function f defined in a domain $D \subset R^2$ and let f is continuous and with respect to y, it satisfies Lipschitz condition in D, that is*

$$|f(x,y_1) - f(x,y_2| \leq \; k \; |y_1 - y_2|, \;\; k > 0,$$

for any two point (x,y_1) and (x,y_2) in D. Then the function $y = \phi(x)$ satisfying the initial value problem $y' = f(x,y)$ and $\phi(x_0) = y_0$, $x\varepsilon I$, is unique.

The following lemma, refer to Gupta and Kirmani [51], presents the sufficient condition which ensures that Lipschitz condition is satisfied in D.

Lemma 2.3.1. *Let f is continuous function in a convex region $D \subset R^2$ and let partial derivative of f that is $\frac{\partial f}{\partial y}$ exists and is continuous in D. Then, f satisfies Lipschitz condition in D.*

Theorem 2.3.2. *Let X be a non-negative continuous random variable with distribution function F(x). Let the past entropy measure of the corresponding n^{th} k-record value $\bar{H}(f_{n,k};t)$ be finite for all $t \geq 0$. Then $\bar{H}(f_{n,k};t)$ characterizes the distribution function uniquely.*

Proof. From (2.2.2), we have

$$\frac{d}{dt}(\bar{H}(f_{n,k};t)) = r_{F_{n,k}}(t) - r_{F_{n,k}}(t)\bar{H}(f_{n,k};t)) - r_{F_{n,k}}(t)\ln r_{F_{n,k}}(t),$$

where $r_{F_{n,k}}(t)$ is the reversed hazard rate function of $X_{n,k}$. Differentiating with respect to t again, we obtain

$$r'_{F_{n,k}}(t) = \frac{\bar{H}'(f_{n,k};t)r_{F_{n,k}}(t) + \bar{H}''(f_{n,k};t)}{-\bar{H}(f_{n,k};t) - \ln r_{F_{n,k}}(t)} \,. \tag{2.3.1}$$

Consider two distribution functions F and F^* such that

$$\bar{H}(f_{n,k};t) = \bar{H}(f^*_{n,k};t) = \gamma(t), \; say.$$

Then $\forall\, t$, from (2.3.1) we get

$$r'_{F_{n,k}}(t) = \psi\left(t, r_{F_{n,k}}(t)\right), \;\; r'_{F^*_{n,k}}(t) = \psi\left(t, r_{F^*_{n,k}}(t)\right),$$

where

$$\psi(t,y) = \left(-\frac{\gamma(t)y+\gamma''(t)}{\gamma(t)+\ln y}\right).$$

Using theorem and lemma by Gupta and Kirmani [51] we get, $r_{F_{n,k}}(t) = r_{F^*_{n,k}}(t), \ \forall\, t$. This proves the uniqueness of reversed hazard rate function. As reversed hazard function characterizes the distribution function uniquely, therefore this characterizes the distribution function $F_{n,k}$ and hence parent distribution uniquely. □

2.4 Some Results On Past Entropy For *k*-Record

2.4.1 A Bound To Past Entropy:

"If $\bar{H}(f_{n,k};t)$ is increasing function of $t>0$, then $r_{n,k}(t) \le \exp(1-\bar{H}(f_{n,k};t))$ and $\bar{H}(f_{n,k};t) \le 1-\ln r_{n,k}(t)$."

Proof. From (2.2.1), we have

$$\bar{H}(f_{n,k};t) = 1 - \frac{1}{F_{n,k}(t)}\int_0^t f_{n,k}(x)\ln r_{n,k}(x)dx.$$

or

$$1-\bar{H}(f_{n,k};t) = \frac{1}{F_{n,k}(t)}\int_0^t f_{n,k}(x)\ln r_{n,k}(x)dx.$$

Now $1-\bar{H}(f_{n,k};t)$ is decreasing function of t as $\bar{H}(f_{n,k};t)$ is increasing, hence for $x \le t$,

$$\begin{aligned} 1-\bar{H}(f_{n,k};t) &\ge \frac{1}{F_{n,k}(t)}\int_0^t f_{n,k}(x)\ln r_{n,k}(t)dx \\ &= \ln r_{n,k}(t). \end{aligned}$$

or

$$r_{n,k}(t) \le \exp(1-\bar{H}(f_{n,k};t)). \tag{2.4.1}$$

Also from (2.4.1), it follows directly

$$\bar{H}(f_{n,k};t) \le 1-\ln r_{n,k}(t). \tag{2.4.2}$$

Hence from (2.4.1) and (2.4.2), we get the result.

2.4.2 Effect Of Monotone Transform:

"For a strictly convex function ϕ with $\phi(0)=0$. Let $Y=\phi(X)$, then

$$\bar{H}(g_{n,k};t)=\bar{H}(f_{n,k};\phi^{-1}(t))+E_{f_{n,k}}[\ln\phi'(X)|X_{n,k}<\phi^{-1}(t)], \tag{2.4.3}$$

where $E_{f_{n,k}}$ denotes the expectation with respect to $f_{n,k}$ and other letters have usual meaning."

Proof. From (2.2.1), the past measure of entropy associated with $Y_{n,k}$ is

$$\begin{aligned}
\bar{H}(g_{n,k};t) &= 1-\frac{1}{G_{n,k}(t)}\int_0^t g_{n,k}(y)\ln r^G_{n,k}(y)dy,\ \ t>0\\
&= 1-\int_0^t \frac{k^n(-\ln\bar{G}(y))^{n-1}(\bar{G}(y))^{k-1}g(y)}{(\Gamma n)G_{n,k}(t)}\ln r^G_{n,k}(y)dy\\
&= 1-\int_0^{\phi^{-1}(t)} \frac{k^n(-\ln\bar{F}(x))^{n-1}(\bar{F}(x))^{k-1}f(x)}{(\Gamma n)F_{n,k}(\phi^{-1}(t))\phi'(x)}\ln\left(\frac{r^F_{n,k}(x)}{\phi'(x)}\right)\phi'(x)dx\\
&= 1-\frac{1}{F_{n,k}(\phi^{-1}(t))}\int_0^{\phi^{-1}(t)} f_{n,k}(x)\ln r^F_{n,k}(x)dx+\frac{1}{F_{n,k}(\phi^{-1}(t))}\int_0^{\phi^{-1}(t)} f_{n,k}(x)\ln\phi'(x)dx\\
&= \bar{H}(f_{n,k};\phi^{-1}(t))+E_{f_{n,k}}[\ln\phi'(X)|X_{n,k}<\phi^{-1}(t)].
\end{aligned}$$

Here $r^F_{n,k}$ and $r^G_{n,k}$ denote the reversed hazard rates of the n^{th} k-record values for the parent distribution functions F and G respectively. □

2.4.3 Stochastic Ordering Of Past Entropy:

Here we prove some order properties for the past entropy measure. First, we give some definitions from Nanda and Paul [87] as follows:

Definition 2.4.1. *A random variable is said to have increasing uncertainty of life (IUL) if $\bar{H}(f;t)$ is increasing in $t\geq 0$.*

Definition 2.4.2. *Let X_1 and X_2 be two random variables denoting the lifetimes of two components with pdfs f_1 and f_2 respectively. Then X_1 is said to be greater than X_2 in past entropy order (written as $X_1 \overset{PE}{\geq} X_2$) if $\bar{H}(f_1;t)\leq\bar{H}(f_2;t),\ \ \forall\ t\geq 0$.*

Example 2.4.1. *If X_1 and X_2 are two exponentially distributed random variable with means a_1 and $2a_1$, then for $n=k=1$, we can easily see that $H(f_{n,k};t)\leq H(g_{n,k};t)$, where $f_{n,k}(x)$ and $g_{n,k}(x)$ represent the pdfs of n^{th} k-record values corresponding to X_1 and X_2 respectively. Therefore X_1 is greater than X_2 in past entropy order that is $X_1 \overset{PE}{\geq} X_2$.*

Next, Theorem 2.4.1 and Theorem 2.4.2 prove results on order properties of the past entropy measure (2.2.1).

Theorem 2.4.1. *Consider two non-negative random variables X_1 and X_2 with $X_{n,k}^{(1)} \stackrel{PE}{\geq} X_{n,k}^{(2)}$ and ϕ be a strictly increasing, differentiable and convex function with $\phi(0) = 0$ and $\phi'(x)$ is continuous with $\phi'(0) \geq 1$. Then $\phi(X_{n,k}^{(1)}) \stackrel{PE}{\geq} \phi(X_{n,k}^{(2)})$, where $X_{n,k}^{(1)}$ and $X_{n,k}^{(2)}$ denote the n^{th} k-record values corresponding to X_1 and X_2 respectively.*

Proof. From (2.4.3), we can write

$$\begin{aligned}\bar{H}(g_{n,k}^{(1)};t) - \bar{H}(g_{n,k}^{(2)};t) &= \bar{H}(f_{n,k}^{(1)};\phi^{-1}(t)) - \bar{H}(f_{n,k}^{(2)};\phi^{-1}(t)) \\ &\quad + E_{f_{n,k}^{(1)}}(\ln\phi'(X)|X_1 < \phi^{-1}(t)) - E_{f_{n,k}^{(2)}}(\ln\phi'(X)|X_2 < \phi^{-1}(t)). \quad (2.4.4)\end{aligned}$$

Here $f_{n,k}^{(1)}$, $f_{n,k}^{(2)}$, $g_{n,k}^{(1)}$ and $g_{n,k}^{(2)}$ represent the pdfs of n^{th} k-record values corresponding to X_1, X_2, $\phi(X_1)$ and $\phi(X_2)$ respectively.
Now, $X_{n,k}^{(1)} \stackrel{PE}{\geq} X_{n,k}^{(2)}$ implies that $\bar{H}(f_{n,k}^{(1)};\phi^{-1}(t)) \leq \bar{H}(f_{n,k}^{(2)};\phi^{-1}(t))$, where as

$$\begin{aligned}&E_{f_{n,k}^{(1)}}(\ln\phi'(X)|X_1 < \phi^{-1}(t)) - E_{f_{n,k}^{(2)}}(\ln\phi'(X)|X_2 < \phi^{-1}(t)) \\ &= \int_0^{\phi^{-1}(t)} \ln\phi'(x)\left(\frac{f_{n,k}^{(1)}(x)}{F_{n,k}^{(1)}(\phi^{-1}(t))} - \frac{f_{n,k}^{(2)}(x)}{F_{n,k}^{(2)}(\phi^{-1}(t))}\right)dx \\ &\leq \ln\phi'(\phi^{-1}(t))\int_0^{\phi^{-1}(t)}\left(\frac{f_{n,k}^{(1)}(x)}{F_{n,k}^{(1)}(\phi^{-1}(t))} - \frac{f_{n,k}^{(2)}(x)}{F_{n,k}^{(2)}(\phi^{-1}(t))}\right)dx \\ &= 0.\end{aligned}$$

The above inequality holds due to $\phi'(x)$ is an increasing function of x. Hence using (2.4.4), we get $\bar{H}(g_{n,k}^{(1)};t) \leq \bar{H}(g_{n,k}^{(2)};t)$, $\forall\, t \geq 0$, which results in $\phi(X_{n,k}^{(1)}) \stackrel{PE}{\geq} \phi(X_{n,k}^{(2)})$. □

Example 2.4.2. *From Example 2.4.1, $X_1 \stackrel{PE}{\geq} X_2$. Let us take $\phi(x) = \lambda x$, where $\lambda \geq 1$. Now $\phi(x)$ is strictly increasing, differentiable and convex function with $\phi(0) = 0$ and $\phi'(x)$ is continuous with $\phi'(0) \geq 1$. Therefore, by Theorem 2.4.1, for $n = k = 1$, $\phi(X_{n,k}^{(1)}) \stackrel{PE}{\geq} \phi(X_{n,k}^{(2)})$.*

Next, we show that the above past entropy order is closed under increasing past entropy measure. First we prove the following lemma.

Lemma 2.4.1. *Consider $Z = aX + b$, where X be any absolutely continuous random vari-*

able, and $a > 0$ and $b \geq 0$ are constants. Then, for $t > b$,

$$\bar{H}(g_{n,k};t) = \bar{H}(f_{n,k};\frac{t-b}{a}) + \ln a, \tag{2.4.5}$$

where $g_{n,k}$ and $f_{n,k}$ are the pdf of n^{th} upper k-record values for Z and X respectively.

Proof. We know that

$$\bar{H}(g_{n,k};t) = \ln G_{n,k}(t) - \frac{1}{G_{n,k}(t)} \int_0^t g_{n,k}(z) \ln g_{n,k}(z) dz.$$

After substituting (1.10.1) and using the transformation $Z = aX + b$ we get,

$$\begin{aligned}\bar{H}(g_{n,k};t) &= \ln F_{n,k}(t) - \frac{1}{F_{n,k}(t)} \int_0^{\frac{t-b}{a}} f_{n,k}(x) \ln\left(\frac{f_{n,k}(x)}{a}\right) dx \\ &= \bar{H}(f_{n,k};\frac{t-b}{a}) + \ln a.\end{aligned}$$

Here $t > b,\ a > 0 \Rightarrow \frac{t-b}{a} > 0$. □

Theorem 2.4.2. *For any two absolutely continuous random variables X and Y, define $Z^x_{n,k} = a_1 X_{n,k} + b_1$ and $Z^y_{n,k} = a_2 Y_{n,k} + b_2$, where $a_1,\ a_2 > 0$ and $b_1,\ b_2 \geq 0$ are constants. Let*
(i) $X_{n,k} \overset{PE}{\geq} Y_{n,k}$, (ii) $a_1 \geq a_2$ and $b_1 \geq b_2$. Then, $Z^x_{n,k} \overset{PE}{\geq} Z^y_{n,k}$ if either $\bar{H}(f_{n,k};t)$ or $\bar{H}(g_{n,k};t)$ is increasing in $t > b_1$. Here $f_{n,k}(x)$ and $g_{n,k}(x)$ denote the p.d.f of n^{th} k-record values corresponding to X and Y respectively.

Proof. Let $\bar{H}(f_{n,k};t)$ is increasing in t. As $\frac{t-b_1}{a_1} \leq \frac{t-b_2}{a_2}$, therefore

$$\bar{H}(f_{n,k};\frac{t-b_1}{a_1}) \leq \bar{H}(f_{n,k};\frac{t-b_2}{a_2}).$$

Also $X_{n,k} \overset{PE}{\geq} Y_{n,k}$

$$\Rightarrow \bar{H}(f_{n,k};\frac{t-b_2}{a_2}) \leq \bar{H}(g_{n,k};\frac{t-b_2}{a_2}).$$

Combining above two inequalities, we have

$$\bar{H}(f_{n,k};\frac{t-b_1}{a_1}) \leq \bar{H}(g_{n,k};\frac{t-b_2}{a_2}). \tag{2.4.6}$$

Using (2.4.5), we get

$$\bar{H}(z^x_{n,k};t) \leq \bar{H}(z^y_{n,k};t).$$

$$\Rightarrow Z^x_{n,k} \overset{PE}{\geq} Z^y_{n,k},$$

this proves the result. Also if $\bar{H}(g_{n,k};t)$ is increasing in t, then proof is on similar lines and hence omitted. □

Corollary 2.4.1. *Let X and Y be two absolutely continuous random variables such that* $X_{n,k} \overset{PE}{\geq} Y_{n,k}$. *Define* $Z^x_{n,k} = aX_{n,k} + b$ *and* $Z^y_{n,k} = aY_{n,k} + b$, *where* $a > 0$ *and* $b \geq 0$ *are constants. Then,* $Z^x_{n,k} \overset{PE}{\geq} Z^y_{n,k}$, *if either* $\bar{H}(f_{n,k};t)$ *or* $\bar{H}(g_{n,k};t)$ *is increasing in* $t > b$.

Although this corollary can be derived directly from the above theorem by taking $a_1 = a_2 = a$ and $b_1 = b_2 = b$, a stronger result stated below as Theorem 2.4.3 can be proved in which condition of increasing entropy measure in the Theorem 2.4.2 has been droped.

Theorem 2.4.3. *Let X and Y be two absolutely continuous random variables. Define* $Z^x_{n,k} = aX_{n,k} + b$ *and* $Z^y_{n,k} = aY_{n,k} + b$, *where* $a > 0$ *and* $b \geq 0$. *Then,* $Z^x_{n,k} \overset{PE}{\geq} Z^y_{n,k}$, *if* $X_{n,k} \overset{PE}{\geq} Y_{n,k}$.

Proof. Proof can be done on the same lines as Theorem 2.4.2 and hence omitted. □

Next, we derive the simplified expression for the computation of past entropy measure for specific distribution functions.

2.5 Past Entropy For Some Specific Distributions

For this first we prove a lemma as given below.

Lemma 2.5.1. *Let* $X_{n,k}$ *denotes the* n^{th} *upper k-record value having pdf* $f_{n,k}(x)$, *and F(x) and f(x) denote the distribution and density function of parent random variable respectively. Then the past entropy measure can be expressed as*

$$\begin{aligned}\bar{H}(f_{n,k};t) =& \ln\left(F_{n,k}(t)\frac{\Gamma n}{k}\right) - \frac{1}{F_{n,k}(t)}\int_0^{-k\ln\bar{F}(t)} \frac{u^{n-1}e^{-u}}{\Gamma n}\ln\left(u^{n-1}f(F^{-1}(1-e^{-\frac{u}{k}}))\right)du \\ &+ \frac{n(k-1)}{k}\frac{F_{n+1,k}(t)}{F_{n,k}(t)}.\end{aligned} \tag{2.5.1}$$

Proof. We have

$$\bar{H}(f_{n,k};t) = \ln F_{n,k}(t) - \frac{1}{F_{n,k}(t)}\int_0^t f_{n,k}(x)\ln f_{n,k}(x)dx.$$

After substituting (1.10.1) and using substitution $-k\ln\bar{F}(x)=u$, we get

$$\bar{H}(f_{n,k};t)=\ln F_{n,k}(t)-\frac{k^n}{F_{n,k}(t)\Gamma n}\int_0^t(-\ln\bar{F}(x))^{n-1}(\bar{F}(x))^{k-1}f(x)$$
$$\ln\left(\frac{k^n}{\Gamma n}(-\ln\bar{F}(x))^{n-1}(\bar{F}(x))^{k-1}f(x)\right)dx.$$
$$=\ln F_{n,k}(t)-\frac{1}{F_{n,k}(t)\Gamma n}\int_0^{-k\ln\bar{F}(t)}u^{n-1}e^{-u}\ln\left(\frac{k}{\Gamma n}u^{n-1}e^{-(1-\frac{1}{k})u}f\left(F^{-1}(1-e^{-\frac{u}{k}})\right)\right)du$$
$$+\frac{n(k-1)}{k}\frac{F_{n+1,k}(t)}{F_{n,k}(t)}$$
$$=\ln F_{n,k}(t)-\ln\frac{k}{\Gamma n}-\frac{1}{F_{n,k}(t)\Gamma n}\int_0^{-k\ln\bar{F}(t)}u^{n-1}e^{-u}\ln\left(u^{n-1}f\left(F^{-1}(1-e^{-\frac{u}{k}})\right)\right)du$$
$$+\frac{n(k-1)}{k}\frac{F_{n+1,k}(t)}{F_{n,k}(t)}$$
$$=\ln\left(F_{n,k}(t)\frac{\Gamma n}{k}\right)-\frac{1}{F_{n,k}(t)}\int_0^{-k\ln\bar{F}(t)}\frac{u^{n-1}e^{-u}}{\Gamma n}\ln\left(u^{n-1}f\left(F^{-1}(1-e^{-\frac{u}{k}})\right)\right)du$$
$$+\frac{n(k-1)}{k}\frac{F_{n+1,k}(t)}{F_{n,k}(t)}.$$

We can also write the (2.5.1) in terms of expectation as

$$\bar{H}(f_{n,k};t)=\ln\left(F_{n,k}(t)\frac{\Gamma n}{k}\right)-\frac{1}{F_{n,k}(t)}.E\left(\ln\left(u^{*n-1}f\left(F^{-1}(1-e^{-\frac{u^*}{k}})\right)\right)\right)$$
$$+\frac{n(k-1)}{k}\frac{F_{n+1,k}(t)}{F_{n,k}(t)}.$$

Here $E(.)$ denotes the expectation and u^* follows the incomplete gamma distribution. Now, the following example obtain the measure of past entropy for n^{th} upper k-record value from exponential distribution in terms of incomplete gamma function which is defined as

$$\Gamma(n;a)=\int_0^a u^{n-1}e^{-u}du \quad a>0.$$

Example 2.5.1. *Let $f^*_{n,k}$ denotes the pdf of n^{th} upper k-record value from standard exponential distribution. Then $F_{n,k}(t)=\frac{\Gamma(n;kat)}{\Gamma(n)}$. Using (2.5.1), we get*

$$H(f^*_{n,k};t)=\ln\frac{\Gamma(n;kat)}{ak}-\frac{(n-1)\Gamma n}{\Gamma(n;kat)}E(\ln\nu^*)+\frac{\Gamma(n+1;kat)}{\Gamma(n;kat)}. \qquad (2.5.2)$$

Here ν^* has the incomplete gamma distribution $\Gamma(n;kat)$.

Remark 2.5.1. *The entropy measure of n^{th} k-record value from distribution F(x) can be* *expressed in terms of past entropy measure from standard exponential distribution as*

$$\bar{H}(f_{n,k};t) = H(f^*_{n,k},t) + \ln a - \frac{1}{k}\frac{\Gamma(n+1;-k\ln\bar{F}(t))}{\Gamma(n;-k\ln\bar{F}(t))} - \frac{\Gamma n}{\Gamma(n;-\ln\bar{F}(t))}. \tag{2.5.3}$$

Proof. The proof follows directly using (2.5.1) and (2.5.2). □

Example 2.5.2. *If X follows the finite range distribution with pdf*

$$f(x) = \frac{a}{b}\left(1-\frac{x}{b}\right)^{a-1}, \quad a>1, \; 0 \le x \le b.$$

and survival function

$$\bar{F}(x) = 1 - F(x) = \left(1-\frac{x}{b}\right)^{a}, \; then$$

$$F_{n,k}(t) = \frac{\Gamma(n;-k\ln\bar{F}(t))}{\Gamma n} = \frac{\Gamma(n;-k\ln\left(1-\frac{t}{b}\right)^{a})}{\Gamma n}.$$

Substituting $-k\ln\bar{F}(x) = u$*, we observe that* $x = F^{-1}(1-e^{-\frac{u}{k}}) = b(1-e^{-\frac{u}{ak}})$.
Using (2.5.1), we get

$$\begin{aligned}
\bar{H}(f_{n,k};t) =& \ln\left(F_{n,k}(t)\frac{\Gamma n}{k}\right) - \frac{1}{F_{n,k}(t)}E\left(\ln \nu^{*n-1}\right)\\
&- \frac{1}{F_{n,k}(t)}\int_0^{-k\ln\bar{F}(t)}\frac{u^{n-1}e^{-u}}{\Gamma n}\ln\left(f\left(F^{-1}(1-e^{-\frac{u}{k}}\right)\right)du\\
&+ \frac{n(k-1)}{k}\frac{F_{n+1,k}(t)}{F_{n,k}(t)}\\
=& \ln\left(F_{n,k}(t)\frac{\Gamma n}{k}\right) - \frac{1}{F_{n,k}(t)}E(\ln\nu^{*n-1}) - \ln\frac{a}{b} + \frac{(a-1)nF_{n+1,k}(t)}{akF_{n,k}(t)}\\
&+ \frac{n(k-1)}{k}\frac{F_{n+1,k}(t)}{F_{n,k}(t)}.
\end{aligned}$$

where ν^* *follows the incomplete gamma distribution* $\Gamma(n,-k\ln\left(1-\frac{t}{b}\right)^{a})$.

Example 2.5.3. *If X follows the Pareto distribution with pdf and cdf as*

$$f(x) = \frac{a}{x^{a+1}}, \; x \ge 1, \; a > 0 \; and \; F(x) = 1 - \frac{1}{x^a}, \; then$$

$$F_{n,k}(t) = \frac{\Gamma(n;ak\ln t)}{\Gamma n} \; and \; f\left(F^{-1}(1-e^{-\frac{u}{k}})\right) = ae^{-\frac{u}{k}(1+\frac{1}{a})}.$$

Using (2.5.1) , we get

$$\bar{H}(f_{n,k};t) = \ln\left(F_{n,k}(t)\frac{\Gamma n}{ak}\right) - \frac{(n-1)}{F_{n,k}(t)}E(\ln\nu^*) + \frac{nF_{n+1,k}(t)}{kF_{n,k}(t)}\left(2+\frac{1}{a}-\frac{1}{k}\right),$$

where ν^* *follows the incomplete gamma distribution* $\Gamma(n; ak\ln t)$.

2.6 Conclusion

Past information measures have found applications in reliability and life testing. In this chapter we have considered a system or a component which was continuously under supervision at regular interval times but at particular instant of time, it was found to be dead. We have defined the measure of uncertainty in past lifetime distribution for such type of systems for k-record values. We have proved that the proposed measure of uncertainty for k-record values determines the underlying distribution uniquely. In addition to a few other properties it satisfies stochastic ordering property. Also we have derived its expression for a few specific distributions.

Chapter 3

Kerridge Measure Of Inaccuracy For Record Statistics

3.1 Introduction

One of the basic problems encountered in reliability theory is the identification of an appropriate probability distribution for lifetime of a component or a system. Various methods like goodness of fit procedures, probability plots etc. are available in literature to find a suitable model followed by the observations. Kerridge inaccuracy [63] provides a useful tool in measuring the two types of errors in expressing the probabilities of various events in performing an experiment. The inaccuracy measure has various applications in different areas of science and technology such as statistical inference, estimation and coding theory (see Nath [83]). Next, the record data evolves in various practical situations like hydrology, sports, industrial stress testing, meteorological analysis and seismology see, for instance, Ahmadi and Arghami [4].

In the preceding chapter we have considered measure of entropy for past lifetime for k-record values. Also we have proved that this entropy measure for past life-

The work reported in this chapter has been published in the paper entitled, **Kerridge Measure of Inaccuracy for Record Statistics** in *Journal of Information and Optimization Sciences*, 2018, 39(5), 1149-1161, and some work has been presented in the International Conference on interdisciplinary Mathematics, Statistics and Computational Techniques held at Manipal University, Jaipur, Dec 22-24, 2016.

time characterizes the distribution function uniquely. If we put $k=1$ in the results for k-record values, we can get the results for ordinary records. In this chapter, we extend the concept of Kerridge measure of inaccuracy to record values and also study some of its properties.
The chapter is organised as follows: In Section 3.2 , we propose an inaccuracy measure between the distribution of n^{th} record values from a sequence of iid random variables and parent distribution and study a characterization result based on this measure. In Section 3.3 , we study the inaccuracy measure for some specific distributions and in Section 3.4 and 3.5 we consider F^{α} distributions and study measure of inaccuracy for some specific F^{α} distributions. Section 3.6 concludes the chapter.

3.2 A Measure Of Inaccuracy For Record Statistics

Corresponding to the inaccuracy measure given by Kerridge [63], we define the measure of inaccuracy between the distribution of n^{th} record value and the parent distribution as

$$H(f_n,f)=-\int_0^{\infty} f_n(x)\log f(x)dx. \tag{3.2.1}$$

Here $f_n(x)$ denotes the pdf of the n^{th} record value given by

$$f_n(x)=\frac{R^{n-1}(x)}{\Gamma(n)}f(x), -\infty<x<\infty, \tag{3.2.2}$$

where $R(x)=-\ln(1-F(x))$.
From (3.2.1), we can write

$$\begin{aligned}
H(f_n,f)=&-\int_0^{\infty} f_n(x)\log\left(\frac{f(x)f_n(x)}{f_n(x)}\right)dx\\
=&\int_0^{\infty} f_n(x)\log\left(\frac{f_n(x)}{f(x)}\right)dx-\int_0^{\infty} f_n(x)\log f_n(x)dx\\
=&H(f_n/f)+H(f_n),
\end{aligned}$$

where $H(f_n/f)$ and $H(f_n)$ denote the Kullback measure of relative information between the distribution of n^{th} record value and the parent distribution and Shannon measure of entropy for the n^{th} record value respectively.

Next, we show that the inaccuracy measure defined above determines the parent distribution function uniquely. To prove this we use the lemma by Goffman and Pedrick [48] which is stated as follows:

Lemma 3.2.1. *A complete orthogonal system for the space* $L_2(0, \infty)$ *is given by the sequence of Laguerre functions*

$$\phi_n(x) = \frac{1}{n!}e^{-\frac{x}{2}}L_n(x),\ n \geq 0.$$

Here $L_n(x)$ *denotes the Laguerre polynomial which is defined as the sum of coefficients of* e^{-x} *in the* n^{th} *derivative of* $x^n e^{-x}$*, that is*

$$L_n(x) = e^x \frac{d^n}{dx^n}(x^n e^{-x}) = \sum_{k=0}^{n}(-1)^k \binom{n}{k} n(n-1)\cdots(k+1)x^k.$$

The completeness of Laguerre functions in $L_2(0, \infty)$ *means that if* $f \in L_2(0, \infty)$ *and* $\int_0^\infty f(x)e^{-\frac{x}{2}}L_n(x)dx = 0,\ \forall n \geq 0$, *then* f *is zero almost everywhere.*

Now, we give the characterization result as follows:

Theorem 3.2.1. *Let* X_1 *and* X_2 *be two random variables with pdfs* $f_1(x)$ *and* $f_2(x)$ *and absolutely continuous cdfs* $F_1(x)$ *and* $F_2(x)$ *respectively. Then* F_1 *and* F_2 *belong to the same family of distributions but for change in location, iff*

$$H(f_{n,1}, f_1) = H(f_{n,2}, f_2),\ n \geq 1. \tag{3.2.3}$$

Here $f_{n,1}(x)$ *and* $f_{n,2}(x)$ *are the density functions of* n^{th} *record value for the parent distributions* $f_1(x)$ *and* $f_2(x)$ *respectively.*

Proof. The necessary part is obviously holds. We need to prove the sufficient part only. For all $n \geq 1$, let

$$H(f_{n,1}(x), f_1(x)) = H(f_{n,2}(x), f_2(x))$$
$$\Rightarrow -\int_0^\infty f_{n,1}(x)\log(f_1(x))dx = -\int_0^\infty f_{n,2}(x)\log(f_2(x))dx$$
$$-\int_0^\infty \frac{(-\log(1-F_1(x)))^{n-1}f_1(x)}{\Gamma(n)}\log(f_1(x))dx = -\int_0^\infty \frac{(-\log(1-F_2(x)))^{n-1}f_2(x)}{\Gamma(n)}\log(f_2(x))dx$$

Putting $u = -\log(1-F_1(x))$ and $u = -\log(1-F_2(x))$, we obtain

$$\int_0^\infty u^{n-1}e^{-u}\log(f_1(F_1^{-1}(1-e^{-u})))du = \int_0^\infty u^{n-1}e^{-u}\log(f_2(F_2^{-1}(1-e^{-u})))du$$

$$\Rightarrow \int_0^\infty \{\log(f_1(F_1^{-1}(1-e^{-u}))) - \log(f_2(F_2^{-1}(1-e^{-u})))\} e^{-u} u^{n-1} du = 0.$$

$$\Rightarrow \int_0^\infty \{\log(f_1(F_1^{-1}(1-e^{-u}))) - \log(f_2(F_2^{-1}(1-e^{-u})))\} e^{\frac{-u}{2}} L_n(u) du = 0.$$

Here $L_n(u)$ is the Laguerre polynomial given in Lemma 3.2.1

Using Lemma 3.2.1, we have

$$f_1(F_1^{-1}(1-e^{-u})) = f_2(F_2^{-1}(1-e^{-u})).$$

Substituting $1-e^{-u} = v$ in the above expression, we get

$$f_1(F_1^{-1}(v)) = f_2(F_2^{-1}(v)), \quad \forall\, v \in (0,1).$$

It is easy to show that $\frac{d(F^{-1}(v))}{dv} = (F^{-1})'(v) = \frac{1}{f(F^{-1}(v))}$. Therefore, we have

$$\begin{aligned} (F_1^{-1})'(v) &= (F_2^{-1})'(v), \ \forall\, v \in (0,1) \\ F_1^{-1}(v) &= F_2^{-1}(v) + a, \end{aligned}$$

where a is constant and this proved the result. □

Next we prove another result to show the effect of monotone transformations on inaccuracy measure defined in (3.2.1) in the following theorem.

Theorem 3.2.2. *Let X be a non-negative and continuous random variable with pdf $f(x)$ and distribution function $F(x)$. Let $Y = \phi(X)$, where ϕ is a strictly monotonically increasing and differentiable function with derivative ϕ', and let G(y) and g(y) denote the distribution and density functions of Y respectively and X_n denotes the n^{th} record value associated with X with pdf f_n and Y_n denotes the n^{th} record value associated with Y with pdf g_n . Then we have*

$$H(g_n, g) = H(f_n, f) - H(f_n, \phi'(x)) \tag{3.2.4}$$

Proof. The probability density function of $Y = \phi(X)$ is $g(y) = \frac{f(\phi^{-\prime}(y))}{\phi'(\phi^{-1}(y))}$. Thus

$$H(g_n, g) = -\int_0^\infty \frac{(-\log G(y))^{n-1}}{\Gamma(n)} g(y) \log g(y) dy.$$

This gives

$$H(g_n, g) = -\int_0^\infty \frac{(-\log F(\phi^{-1}(y)))^{n-1}}{\Gamma(n)} \frac{f(\phi^{-1}(y))}{\phi'(\phi^{-1}(y))} \log \frac{f(\phi^{-1}(y))}{\phi'(\phi^{-1}(y))} dy.$$

By taking $x = \phi^{-1}(y)$, we obtain

$$H(g_n, g) = -\int_0^\infty \frac{(-\log F(x))^{n-1}}{\Gamma(n)} f(x) \log \frac{f(x)}{\phi'(x)} dx$$

$$= -\int_0^\infty \frac{(-\log F(x))^{n-1}}{\Gamma(n)} f(x) \log f(x) dx + \int_0^\infty \frac{(-\log F(x))^{n-1}}{\Gamma(n)} f(x) \log \phi'(x) dx.$$

This can be written as

$$H(g_n, g) = H(f_n, f) - H(f_n, \phi^{'}(x)).$$

This proves the result. □

Remark 3.2.1. *Let $Y = aX + b$, where X be any absolutely continuous random variable and $a > 0,\; b$ are constants. Then*

$$\phi^{'}(x) = a$$

$$H(g_n, g) = H(f_n, f) + \log a\ .$$

Thus inaccuracy measures defined in (3.2.1) is invariant under location but not under scale transformation.

In particular, if $a = 1$, that is $Y = X + b$, then $H(g_n, g) = H(f_n, f)$, where f and g denote the p.d.f for X and Y respectively.

3.3 Meaure Of Inaccuracy For Some Specific Distributions

Following are the expressions for the proposed inaccuracy measure (3.2.1) for some specific probability distributions for a random variable X.

(i) Uniform Distribution

Consider a random variable X having uniform distribution over (a, b), $a \geq 0$, $a < b$, then pdf and cdf of X is

$$f(x) = \frac{1}{b-a} \quad a \leq x \leq b \quad and \quad \bar{F}(x) = \frac{b-x}{b-a}$$

Then from (3.2.1),

$$\begin{aligned} H(f_n,f) &= -\int_0^\infty f_n(x)\log f(x)dx \\ &= \frac{(-1)^{n+1}\ln(b-a)}{(b-a)(n-1)!}\int_a^b \left(\ln\frac{b-x}{b-a}\right)^{n-1} dx \\ &= \log(b-a). \end{aligned} \tag{3.3.1}$$

Let $\triangle_n = H(f_{n+1},f) - H(f_n,f)$ be the n^{th} inaccuracy differential, that is , the change in inaccuracy in observing the record value from the n^{th} to the $(n+1)^{th}$. In case of uniform distribution $\triangle_n = H(f_{n+1},f) - H(f_n,f) = 0$, that is , measure of inaccuracy $H(f_n,f)$ remains constant for all n for uniform distribution.

(ii) Exponential Distribution

Consider a random variable X having exponential distribution with parameter $a > 0$, then pdf and cdf is given by

$$f(x) = ae^{-ax} \quad and \quad \bar{F}(x) = e^{-ax},$$

Substituting this in (3.2.1) we get

$$\begin{aligned} H(f_n,f) &= -\frac{1}{(n-1)!}\int_0^\infty (ax)^{(n-1)} ae^{-ax}\log(ae^{-ax})dx \\ &= -\frac{a^n}{(n-1)!}\int_0^\infty x^{(n-1)}e^{-ax}\ln(ae^{-ax})dx. \\ &= -\frac{a^n}{(n-1)!}\int_0^\infty x^{(n-1)}e^{-ax}(\ln a - ax)\,dx. \\ &= -\log a + n. \end{aligned} \tag{3.3.2}$$

We observe that for a fixed value of n, inaccuracy of n^{th} record value for exponential distribution decreases with increasing value of the parameter $a > 0$.
Similarly, if a is fixed then $H(f_n,f)$ increases with increase in sample size. In this case we obtain $\Delta_n = H(f_{n+1},f) - H(f_n,f) = 1, \forall\, n$, that is n^{th} inaccuracy differential does not depend on the sample size.

(iii) Weibull Distribution

For Weibull distribution pdf is

$$f(x) = abx^{b-1}\exp\left\{-ax^b\right\},\ \ a,\ b > 0,\ x > \ 0$$

where a and b are parameters. The survival function is

$$\bar{F}(x) = e^{-ax^b}\ .$$

We have

$$H(f_n, f) = -\int_0^\infty \frac{abx^{b-1}e^{-ax^b}\log(abx^{b-1}e^{-ax^b})(-\log(e^{-ax^b}))^{n-1}}{(n-1)!}dx$$

Substituting $-\log\bar{F}(x) = -\log(e^{-ax^b}) = t$, we obtain

$$\begin{aligned} H(f_n, f) \ &= \ -\frac{b-1}{b}\int_0^\infty \frac{t^{n-1}e^{-t}\log t}{\Gamma(n)}dt \ - \ \frac{1}{b}\int_0^\infty \frac{t^{n-1}e^{-t}\log(ab^b)}{\Gamma(n)}dt \ + \ \int_0^\infty \frac{t^n e^{-t}}{\Gamma(n)}dt. \\ &= -\frac{b-1}{b}\Psi(n) - \frac{\log a}{b} - \log b + n. \end{aligned} \tag{3.3.3}$$

In particular for $b = 1$, (3.3.3) reduces to the n^{th} record inaccuracy for exponential distribution.

The n^{th} inaccuracy differential is

$$\begin{aligned} \Delta_n &= H(f_{n+1}, f) - H(f_n, f) \\ &= \frac{b-1}{b}\left(\Psi(n) - \Psi(n+1)\right) + 1 \\ &= \left(\frac{1-b}{bn}\right) + 1 \end{aligned}$$

3.4 F^α Distributions

F^α distributions are quite important concept in statistics because of their utility in modelling and analysis of lifetime data. Various classes of F^α distributions have been developed in literature by various researchers. Let X be an absolutely continuous positive random variable. Then X is said to have F^α distribution if its cumulative distribution function is given by $G(x) = \ F^\alpha(x) = \ [F(x)]^\alpha,\ \alpha > 0,$ the α-th power of the baseline distribution function $F(x)$, refer to Ahsanullah et al. [8]. The distribution $G(x)$ is also called an exponentiated distribution of given baseline distribution function $F(x)$. Its probability density function is given by $g(x)=\alpha f(x)F^{\alpha-1}(x),\ \alpha>0,$ where $f(x)=\frac{dF(x)}{dx}$ is the probability density function of the random variable X.

Further we know that, two random variables X and Y satisfy the proportional reversed hazard rate model (refer Gupta et.al. [49]) with proportionality constant $\alpha\ (>0)$, if

$$G(x)=F^{\alpha}(x) \tag{3.4.1}$$

Here $F(x)$ is the baseline distribution and $G(x)$ can be considered as some reference distribution.

Next, the reverse hazard rate of $G(x)$ denoted by $\lambda_G(x)$ is given as $\lambda_G(x)=\frac{g(x)}{G(x)}$. Substituting values of pdf $g(x)$ and $G(x)$ from above, we get $\lambda_G(x)=\frac{\alpha f(x)F^{\alpha-1}(x)}{F^{\alpha}(x)}=\alpha\lambda_F(x)$, where $\lambda_F(x)$ denotes the reverse hazard rate corresponding to the distribution $F(x)$. Thus in case of power distributions, that reverse hazard rate function of $G(x)$ is proportional to reverse hazard rate function of $F(x)$ with proportionality constant α.

3.5 Measure of Inaccuracy For F^{α} Distributions

In this section we introduce the inaccuracy measure for n^{th} lower record value for some of the F^{α} distributions.

(i) Gompertz-Verhulst Exponentiated Distribution

Gompertz-Verhulst Exponentiated distribution is used to compare known human mortality tables and to represent population growth defined by the cdf and pdf given as follows

$$G(x)=(1-\rho e^{-\lambda x})^{\alpha}$$

and

$$g(x)=\alpha(1-\rho e^{-\lambda x})^{\alpha-1}(\rho\lambda e^{-\lambda x}),$$

where $x>\frac{1}{\lambda}\ln\rho>0$ and $\rho,\lambda,\alpha>0$. Then

$$H(g_n,g)=-\int_0^{\infty}\frac{(-\ln(1-\rho e^{-\lambda x})^{\alpha})^{n-1}\alpha(1-\rho e^{-\lambda x})^{\alpha-1}\rho\lambda e^{-\lambda x}\ln(\alpha(1-\rho e^{-\lambda x})^{\alpha-1}\rho\lambda e^{-\lambda x}))}{\Gamma n}dx.$$

Substituting $-\ln(1-\rho e^{-\lambda x})^{\alpha}=t$, we will get

$$H(g_n,g)=\frac{(\alpha-1)n}{\alpha}-\ln\alpha\lambda-\int_0^{\infty}\frac{e^{-t}t^{n-1}\ln(1-e^{-\frac{t}{\alpha}})}{\Gamma n}dt.$$

Using $-\ln(1-e^{-\frac{t}{\alpha}})=\sum_{j=1}^{\infty}\frac{e^{-\frac{jt}{\alpha}}}{j}$, we obtain

$$H(g_n,g)=\frac{(\alpha-1)n}{\alpha}-\ln\alpha\lambda+\frac{1}{\Gamma(n)}\sum_{j=1}^{\infty}\frac{1}{j}\int_0^{\infty}e^{-(1+\frac{j}{\alpha})t}t^{n-1}dt$$

Let $(1+\frac{j}{\alpha})t=\theta$ then

$$H(g_n,g)=\frac{(\alpha-1)n}{\alpha}-\ln\alpha\lambda+\frac{1}{\Gamma(n)}\sum_{j=1}^{\infty}\frac{1}{j}\int_0^{\infty}\frac{e^{-\theta}\theta^{n-1}d\theta}{(1+\frac{j}{\alpha})^n}$$

This gives

$$H(g_n,g)=\frac{(\alpha-1)n}{\alpha}-\ln\alpha\lambda+\sum_{j=1}^{\infty}\frac{1}{j(1+\frac{j}{\alpha})^n} \tag{3.5.1}$$

$$H(g_1,g)=\frac{(\alpha-1)}{\alpha}-\ln\alpha\lambda+\sum_{j=1}^{\infty}\frac{1}{j(1+\frac{j}{\alpha})}$$

$$\Delta_n=H(g_{n+1},g)-H(g_n,g)=\frac{\alpha-1}{\alpha}-\sum_{j=1}^{\infty}\frac{1}{\alpha(1+\frac{j}{\alpha})^{n+1}}.$$

(ii) Power Function Distribution

A random variable X is said to have power function distribution if its cdf and pdf are given by

$$G(x)=\left(\frac{x}{\lambda}\right)^{\alpha},\ \ 0<x<\lambda,$$

and

$$g(x)=\frac{\alpha}{\lambda}\left(\frac{x}{\lambda}\right)^{\alpha-1},\ \ \lambda>0,\ \ \alpha>0.$$

Using the substitution $-\ln(\frac{x}{\lambda})^{\alpha}=t$, we get

$$\begin{aligned}H(g_n,g)&=-\int_0^{\infty}\frac{t^{n-1}e^{-t}(\ln(\frac{\alpha}{\lambda})+\ln(e^{-\frac{t}{\alpha}})^{\alpha-1})}{\Gamma(n)}dt\\&=\ln\left(\frac{\lambda}{\alpha}\right)+\frac{(\alpha-1)n}{\alpha}.\end{aligned}$$

Now put $n=1$,

$$H(g_1,g)=\ln(\frac{\lambda}{\alpha})+\frac{(\alpha-1)}{\alpha}.$$

Also

$$\Delta_n=H(g_{n+1},g)-H(g_n,g)=\frac{\alpha-1}{\alpha},$$

which means that difference between inaccuracy measures of two consecutive record values from power function distribution does not depends on n. For $\alpha=1$ it is reduced to uniform distribution.

(iii) Generalized Exponential Or Exponentiated Exponential Distribution

The exponentiated exponential distribution was introduced by Gupta and Kundu [50]. A random variable X is said to have exponentiated exponential distribution if its probability density function (pdf) and cumulative distribution function (cdf) are given by

$$f(x)=\alpha\lambda e^{-\lambda x}(1-e^{-\lambda x})^{\alpha-1},\ x>0 \tag{3.5.2}$$

and

$$F(x)=\{1-\exp(-\lambda x)\}^{\alpha},\ \alpha>0,\ \lambda>0, \tag{3.5.3}$$

respectively, which is α^{th} power of cdf of standard exponential distribution. Therefore for $\alpha=1$, (3.5.2) is exponential .
This distribution has some important physical significance. We know that a parallel system consisting of α components, works only when at least one of the α-components works. If the lifetime of the components are independent and identically distributed random variables and follows the exponential distribution, then lifetime distribution of the system can be defined as (3.5.2). This is the particular case of Gompertz-Verhulst Exponentiated distribution, when $\rho=1$. Putting $\rho=1$ in (3.5.1), we get $H(g_n,g)$ same as that of Gompertz-Verhulst Exponentiated distribution.

(iv) Pareto Type 2 Distribution

The cdf and pdf of this distribution is given by

$$G(x) = \left(1 - \left(\frac{1}{1+x}\right)^{\beta}\right)^{\alpha}, \quad \beta,\ \alpha > 0,$$

and

$$g(x) = \alpha\beta\left(1 - \left(\frac{1}{1+x}\right)^{\beta}\right)^{\alpha-1}\left(\frac{1}{1+x}\right)^{\beta+1}, \quad x > 0.$$

Now, the inaccuracy measure comes out to be

$$H(g_n, g) = -\frac{1}{\Gamma(n)}\int_0^{\infty}\left\{-\ln\left(1-(\frac{1}{1+x})^{\beta}\right)^{\alpha}\right\}^{n-1} g(x)\ln\alpha\beta(1-(\frac{1}{1+x})^{\beta})^{\alpha-1}(\frac{1}{1+x})^{\beta+1})dx.$$

By putting $\left(1-(\frac{1}{1+x})^{\beta}\right)^{\alpha} = t$, we get

$$H(g_n, g) = -\int_0^{\infty}\frac{t^{n-1}e^{-t}\ln\left(\alpha\beta e^{\frac{-t(\alpha-1)}{\alpha}}(1-e^{\frac{-t}{\alpha}})^{\frac{\beta+1}{\beta}}\right)}{\Gamma(n)}dt$$

Using $-\ln(1-e^{-\frac{t}{\alpha}}) = \sum_{j=1}^{\infty}\frac{e^{-\frac{jt}{\alpha}}}{j}$, we obtain

$$= -\ln(\alpha\beta) + \frac{n(\alpha-1)}{\alpha} + \frac{(\beta+1)}{\beta\Gamma(n)}\sum_{j=1}^{\infty}\frac{1}{j}\int_0^{\infty}t^{n-1}e^{-t(1+\frac{j}{\alpha})}dt.$$

Let $(1+\frac{j}{\alpha})t = \theta$, then

$$H(g_n, g) = -\ln(\alpha\beta) + \frac{n(\alpha-1)}{\alpha} + \frac{(\beta+1)}{\beta}\sum_{j=1}^{\infty}\frac{1}{j(1+\frac{j}{\alpha})^n}\,. \qquad (3.5.4)$$

$$H(g_1, g) = -\ln(\alpha\beta) + \frac{(\alpha-1)}{\alpha} + \frac{(\beta+1)}{\beta}\sum_{j=1}^{\infty}\frac{1}{j(1+\frac{j}{\alpha})}\,.$$

$$\Delta_n = H(g_{n+1}, g) - H(g_n, g) = \frac{(\alpha-1)}{\alpha} - \frac{(\beta+1)}{\beta}\sum_{j=1}^{\infty}\frac{1}{\alpha(1+\frac{j}{\alpha})^{n+1}}\,.$$

When $\alpha = 1$, (3.5.4) gives the inaccuracy measure for Standard Pareto distribution. Putting $\alpha = 1$ in (3.5.4), we get

$$H(g_n, g) = -\ln(\beta) + \frac{(\beta+1)}{\beta}\sum_{j=1}^{\infty}\frac{1}{j(1+j)^n}\,.$$

3.6 Conclusion

We have extended the concept of Kerridge measure of inaccuracy for record values by defining the inaccuracy measure between the distribution of n^{th} record value and the parent distribution. Also we have proved that the proposed measure characterizes the parent distribution uniquely. Some properties of the proposed measure like effect of monotone transformation have been discussed. The measure has been studied for some specific distributions. Also keeping in mind the usefulness of F^{α} distributions, the inaccuracy measure has been discussed for that class of distributions also.

Chapter 4

Measure Of Inaccuracy And Residual Inaccuracy Measure For k-Records

---

4.1 Introduction

The inaccuracy measure as given by Kerridge [63] is defined as

$$H(f,g) = -\int_0^\infty f(x)\log g(x)dx. \tag{4.1.1}$$

Here $f(x)$ is the actual distribution and $g(x)$ is the predicted one. When $g(x) = f(x)$ for all x, (4.1.1) becomes the Shannon's entropy.
The measure of information and inaccuracy are associated as $H(f,g) = H(f) + H(f/g)$, where $H(f/g)$ represents the Kullback-Leibler [67] relative information measure of X about Y, defined as

$$H(f/g) = \int_0^\infty f(x)\log\frac{f(x)}{g(x)}dx. \tag{4.1.2}$$

In the previous chapter we have discussed the inaccuracy measure between the sequence of record values and the parent distribution. Recently the concept of

---

The results reported in this chapter have been published in the paper entitled **Measure of Inaccuracy and k-Record Statistics** in *Bulletin of Calcutta Mathematical Society*, 2018, 110 (2), 151-166, and some work has been presented in National Seminar on Recent Developments in Mathematical Sciences held at MDU, Rohtak, Mar 07-08, 2017.

k-record values has found great importance in the various fields as the record values does not generate frequently, so it makes statistical inference based on records very difficult to perform. To overcome these difficulties Dziubdziela and Kopocinski [35] introduced the model of k-record statistics. Instead of observing the sequence of largest values, he observed the sequence of k^{th} largest values.
Statistical inference problems based on k-records have been considered by several authors, see, Berred [22], Malinowska and Szynal [78], Ahmadi and Doostparast [6] and Mary and Chacko [79]. So it is worthwhile, to extend the concept of measure of inaccuracy as given by Kerridge [63] and dynamic measure of inaccuracy between two residual lifetimes as given by Taneja et al. [117] to k-record values. In the present chapter we propose and study a measure of inaccuracy between n^{th} and m^{th} upper k-record values and also study a residual measure of inaccuracy between the n^{th} upper k-record value and the parent distribution.
The chapter is organised as follows : In Section 4.2, the inaccuracy measure between distributions of n^{th} and m^{th} upper k-records is proposed. A general expression for this inaccuracy measure is given in the Section 4.3 and the inaccuracy measure between k-record values associated with Uniform, Exponential, Weibull, Pareto and Finite range distributions are presented . Section 4.4 is devoted to the characterization result. In Section 4.5 we have given an expression for residual measure of inaccuracy and a characterization for it. The chapter ends with conclusion.

4.2 Measure Of Inaccuracy Between Distributions Of n^{th} And m^{th} Upper k-Records

The Shannon entropy for the n^{th} upper k-record values is given as

$$H(f_{n,k}) = -\int_0^\infty f_{n,k}(x)\ln f_{n,k}(x)dx, \tag{4.2.1}$$

where $f_{n,k}(x)$ is the p.d.f of n^{th} upper k-record value as given by (1.10.1). Also Kullback measure of relative information between n^{th} and m^{th} upper k-records is given as

$$H(f_{n,k}/f_{m,k}) = \int_0^\infty f_{n,k}(x)\ln\left(\frac{f_{n,k}(x)}{f_{m,k}(x)}\right)dx, \tag{4.2.2}$$

refer to, Ahmadi and Mohtashami [7]. Adding (4.2.1) and (4.2.2), we get

$$H(f_{n,k}) + H(f_{n,k}/f_{m,k}) = -\int_0^\infty f_{n,k}(x)\ln f_{m,k}(x)dx = H(f_{n,k}, f_{m,k}), \tag{4.2.3}$$

which can be considered as the inaccuracy measure between n^{th} and m^{th} upper k-records. In particular if $m = 1$ and $k = 1$ then (4.2.3) gives the inaccuracy measure between n^{th} k-record and the parent distribution refer to, (3.2.1).

4.3 Measure Of Inaccuracy For k-Record Value Obtained For Some Specific Distributions

First we prove the following result.

Lemma 4.3.1. *The inaccuracy measure $H(f_{n,k}, f_{m,k})$ between n^{th} and m^{th} upper k-record values can be expressed as*

$$\begin{aligned} H(f_{n,k}, f_{m,k}) = -\ln\frac{k^m}{\Gamma m} - (m-1)\Psi(n) + (m-1)\ln k + \frac{n}{k}(k-1) - \frac{k^n}{\Gamma n}. \\ \int_0^\infty u^{n-1}e^{-uk}\ln f(F^{-1}(1-e^{-u}))du. \end{aligned} \tag{4.3.1}$$

Proof. The inaccuracy measure (4.2.3) is

$$H(f_{n,k}, f_{m,k}) = -\int_0^\infty f_{n,k}(x)\ln f_{m,k}(x)dx,\ \forall\, n, m \geq 1\ .$$

Using (1.10.1), this becomes

$$H(f_{n,k}, f_{m,k}) = -\int_0^\infty \frac{k^n}{\Gamma n}(-\ln\bar{F}(x))^{n-1}(\bar{F}(x))^{k-1}f(x)\ln\left(\frac{k^m}{\Gamma m}(-\ln\bar{F}(x))^{m-1}(\bar{F}(x))^{k-1}f(x)\right)dx.$$

Substituting $-\ln\bar{F}(x) = u$ and hence $x = F^{-1}(1-e^{-u})$, we have

$$\begin{aligned} H(f_{n,k}, f_{m,k}) &= -\int_0^\infty \frac{k^n}{\Gamma n}u^{n-1}e^{-ku}\ln\left(\frac{k^m}{\Gamma m}u^{m-1}e^{-u(k-1)}f(F^{-1}(1-e^{-u})\right)du \\ &= -\frac{k^n}{\Gamma n}\ln\frac{k^m}{\Gamma m}\int_0^\infty u^{n-1}e^{-ku}du - \frac{k^n}{\Gamma n}(m-1)\int_0^\infty u^{n-1}e^{-ku}\ln u du \\ &+ \frac{k^n}{\Gamma n}(k-1)\int_0^\infty u^n e^{-ku}du - \frac{k^n}{\Gamma n}\int_0^\infty u^{n-1}e^{-ku}\ln f(F^{-1}(1-e^{-u})du. \end{aligned}$$

Using Gamma function as $\Gamma n = \int_0^\infty u^{n-1}e^{-u}du$ and after some simplifications, we get

$$H(f_{n,k}, f_{m,k}) = -\ln\frac{k^m}{\Gamma m} - (m-1)\Psi(n) + (m-1)\ln k + \frac{n}{k}(k-1) - \frac{k^n}{\Gamma n}\int_0^\infty u^{n-1}e^{-uk}\ln f(F^{-1}(1-e^{-u}))du\,,$$

this proves the result. □

Here $\Psi(x)$ denotes digamma function which is logarithmic derivative of the gamma function, given by

$$\Psi(x) = \frac{d}{dx}\{\ln\Gamma(x)\} = \frac{\Gamma'(x)}{\Gamma(x)}.$$

Next, using (4.3.1), we obtain the inaccuracy measure $H(f_{n,k}, f_{m,k})$ for certain specific distributions.

1. Exponential Distribution

Let X be a random variable having the exponential distribution over $(0,\infty)$, then its density and distribution functions are given respectively by

$$f(x) = \theta e^{-\theta x} \text{ ,and } \bar{F}(x) = 1 - F(x) = e^{-\theta x}.$$

Substituting $-\ln\bar{F}(x) = u$, we observe that $x = F^{-1}(1-e^{-u}) = \left(\frac{u}{\theta}\right)$ and for computing $H(f_{n,k}, f_{m,k})$, we have

$$f(F^{-1}(1-e^{-u})) = f(\frac{u}{\theta}) = \theta e^{-u}\ .$$

From (4.3.1)

$$\begin{aligned} H(f_{n,k}, f_{m,k}) &= -\ln\frac{k^m}{\Gamma m} - (m-1)\Psi(n) + (m-1)\ln k + \frac{n}{k}(k-1) - \frac{k^n}{\Gamma n}\int_0^\infty u^{n-1}e^{-uk}\ln(\theta e^{-u})du \\ &= -\ln\frac{k^m}{\Gamma m} - (m-1)\Psi(n) + (m-1)\ln k + \frac{n}{k}(k-1) - \frac{k^n}{\Gamma n}\ln\theta\int_0^\infty u^{n-1}e^{-uk}du \\ &+ \frac{k^n}{\Gamma n}\int_0^\infty u^n e^{-uk}du. \end{aligned}$$

After substituting $ku = t$ and some simplifications we get

$$H(f_{n,k}, f_{m,k}) = -\ln\frac{k^m}{\Gamma m} - (m-1)\Psi(n) + (m-1)\ln k + n - \ln\theta \ . \qquad (4.3.2)$$

When $k = 1$, (4.3.2) reduces to

$$H(f_n, f_m) = \ln\Gamma m - (m-1)\Psi(n) + n - \ln\theta,$$

the inaccuracy measure between n^{th} and m^{th} upper record values. When $m = k = 1$, (4.3.2) reduces to

$$H(f_n, f) = n - \ln\theta,$$

the inaccuracy measure between n^{th} upper record values and parent random variable refer to, (3.3.2).

When $n = m$ and $k = 1$ (4.3.2) reduces to

$$H(f_n) = log\Gamma n - (n-1)\Psi(n) + n - log\theta, \qquad (4.3.3)$$

the shannon entropy of n^{th} record values for exponential variate.
Also when $n = m = k = 1$ this comes out to be

$$H(f) = 1 - \ln\theta,$$

the shannon entropy of parent distribution for exponential variate.

2. Uniform Distribution

If a random variable X is uniformly distributed over $(a\ ,b),\ \ a < b$, then its density and distribution functions are given respectively by

$$f(x) = \frac{1}{b-a} \text{ and } F(x) = \frac{x-a}{b-a}\ ,\ \ a < x < b.$$

From (4.3.1)

$$H(f_{n,k}, f_{m,k}) = -\ln\frac{k^m}{\Gamma m} - (m-1)\Psi(n) + (m-1)\ln k + \frac{n}{k}(k-1) - \frac{k^n}{\Gamma n}\int_0^\infty u^{n-1}e^{-uk}\ln f(x)du.$$

Thus inaccuracy measure between n^{th} and m^{th} k-record value for uniform distribution is given as

$$H(f_{n,k},f_{m,k}) = -\ln\frac{k^m}{\Gamma m} - (m-1)\Psi(n) + (m-1)\ln k + \frac{n}{k}(k-1) - \frac{k^n}{\Gamma n}\int_0^\infty u^{n-1}e^{-uk}\ln\frac{1}{b-a}du.$$

This gives

$$H(f_{n,k},f_{m,k}) = -\ln\frac{k^m}{\Gamma m} - (m-1)\Psi(n) + (m-1)\ln k + \frac{n}{k}(k-1) + \ln(b-a). \quad (4.3.4)$$

When $n = m$ and $k = 1$ (4.3.4) reduces to

$$H(f_n) = log(\Gamma n) - (n-1)\Psi(n) + log(b-a) \quad (4.3.5)$$

the Shannon entropy of n^{th} record values for uniform variate.
When $m = k = 1$ (4.3.4) comes out to be

$$H(f_n, f) = \ln(b-a),$$

the inaccuracy measure between n^{th} upper record values and parent distribution, refer to (3.3.1) .

Also when $n = m = k = 1$ this comes out to be

$$H(f) = \ln(b-a),$$

the Shannon entropy of parent distribution for uniform variate.

3. Pareto Distribution

Let X be a random variable having Pareto distribution with pdf

$$f(x) = \frac{a}{x^{a+1}},\ x \geq 1,\ a > 0.$$

Substituting $-\ln\bar{F}(x) = u$, we observe that $x = F^{-1}(1-e^{-u}) = e^{\frac{u}{a}}$ and for computing $H(f_{n,k},f_{m,k})$, we have

$$f(F^{-1}(1-e^{-u})) = ae^{-(u+\frac{u}{a})}\ .$$

Thus using (4.3.1) inaccuracy measure (4.2.3) between n^{th} and m^{th} upper k-record values for Pareto distribution is given as

$$H(f_{n,k},f_{m,k}) = -\ln\frac{k^m}{\Gamma m} - (m-1)\Psi(n) + (m-1)\ln k + \frac{n}{k}(k-1) - \frac{k^n}{\Gamma n}\int_0^\infty u^{n-1}e^{-uk}\ln(ae^{-(u+\frac{u}{a})})du$$
$$= -\ln\frac{k^m}{\Gamma m} - (m-1)\Psi(n) + (m-1)\ln k + \frac{n}{k}(k-1) - \frac{k^n}{\Gamma n}\int_0^\infty u^{n-1}e^{-uk}\ln a du +$$
$$\frac{k^n}{\Gamma n}(1+\frac{1}{a})\int_0^\infty u^n e^{-uk}du.$$

This gives

$$H(f_{n,k},f_{m,k}) = \ln\frac{\Gamma m}{k} - (m-1)\Psi(n) + n - \ln a + \frac{n}{ak}. \quad (4.3.6)$$

When $k = 1$, (4.3.6) reduces to

$$H(f_n,f_m) = \ln(\Gamma m) - (m-1)\Psi(n) + n - \ln a + \frac{n}{a},$$

When $m = k = 1$, (4.3.6) reduces to

$$H(f_n,f) = n - \ln a + \frac{n}{a},$$

the inaccuracy measure between n^{th} and m^{th} record values for pareto distribution.

When $n = m$ and $k = 1$ (4.3.6) reduces to

$$H(f_n) = log(\Gamma n) - (n-1)\Psi(n) + n - loga + \frac{n}{a}$$

the Shannon entropy of n^{th} record values for Pareto variate.
Also when $n = m = k = 1$, (4.3.6) comes out to be

$$H(f) = 1 - \ln a + \frac{1}{a},$$

the Shannon entropy of parent distribution for Pareto variate.

4. Weibull Distribution

The pdf of Weibull distribution is

$$f(x) = abx^{b-1}\exp\left(-ax^b\right),\ \ a,\ b > 0,\ x > \ 0,$$

where a and b are scale and shape parameters respectively. The survival function is

$$\bar{F}(x) = 1 - F(x) = e^{-ax^b}.$$

Substituting $-\ln \bar{F}(x) = u$, we observe that $x = F^{-1}(1-e^{-u}) = (\frac{u}{a})^{\frac{1}{b}}$ and for computing $H(f_{n,k}, f_{m,k})$, we have

$$f(F^{-1}(1-e^{-u})) = \left(ba^{\frac{1}{b}}\right)(u)^{\frac{(b-1)}{b}} e^{-u}.$$

From (4.3.1)

$$H(f_{n,k}, f_{m,k}) = -\ln \frac{k^m}{\Gamma m} - (m-1)\Psi(n) + (m-1)\ln k + n - \frac{n}{k} - \frac{k^n}{\Gamma n}\int_0^\infty u^{n-1}e^{-uk}\ln\left(ba^{\frac{1}{b}}u^{\frac{(b-1)}{b}}e^{-u}\right)du$$

$$= \ln(\frac{\Gamma m}{k}) - (m-1)\Psi(n) + n - \frac{n}{k} - \frac{k^n}{\Gamma n}\ln\left(ba^{\frac{1}{b}}\right)\int_0^\infty u^{n-1}e^{-uk}du$$
$$-\frac{k^n(b-1)}{b\Gamma n}\int_0^\infty u^{n-1}e^{-uk}\ln u du + \frac{k^n}{\Gamma n}\int_0^\infty u^n e^{-uk}du.$$

After some simplifications, we get

$$H(f_{n,k}, f_{m,k}) = \ln\left(\frac{\Gamma m}{k^{\frac{1}{b}}}\right) - (m - \frac{1}{b})\Psi(n) + n - \ln(ba^{\frac{1}{b}}). \quad (4.3.7)$$

If $k = 1$, this gives

$$H(f_n, f_m) = \ln(\Gamma m) - (m - \frac{1}{b})\Psi(n) + n - \ln(ba^{\frac{1}{b}}),$$

this is the inaccuracy measure between n^{th} and m^{th} record values. If $m = k = 1$, (4.3.7) gives

$$H(f_n, f) = -(1 - \frac{1}{b})\Psi(n) + n - \ln(ba^{\frac{1}{b}}),$$

the inaccuracy measure between n^{th} upper record values and parent random variable, refer to (3.3.3).

If $n = m$ and $k = 1$, then (4.3.7) gives

$$H(f_n) = \ln(\Gamma n) - (n - \frac{1}{b})\Psi(n) + n - \ln g(ba^{\frac{1}{b}}),$$

the Shannon entropy for n^{th} record values.

If $n=m=k=1$, then (4.3.7) becomes

$$H(f) = -(1-\frac{1}{b})\Psi(1)+1-\ln(ba^{\frac{1}{b}}),$$

the shannon entropy of parent distribution for weibull variate.

5. Finite Range Distribution

The pdf of the finite range distribution is given by

$$f(x) = \frac{a}{b}\left(1-\frac{x}{b}\right)^{a-1}, \quad a>1, \ 0 \le x \le b.$$

The survival function is

$$\bar{F}(x) = 1-F(x) = \left(1-\frac{x}{b}\right)^{a}.$$

Substituting $-\ln\bar{F}(x)=u$, we observe that $x=F^{-1}(1-e^{-u}) = b(1-e^{-\frac{u}{a}})$ and for computing $H(f_{n,k}, f_{m,k})$, we have

$$f(F^{-1}(1-e^{-u})) = \frac{ae^{-u+\frac{u}{a}}}{b}.$$

From (4.3.1) gives

$$H(f_{n,k}, f_{m,k}) = -\ln\frac{k^m}{\Gamma m} - (m-1)\Psi(n) + (m-1)\ln k + \frac{n}{k}(k-1) - \frac{k^n}{\Gamma n}\int_0^\infty u^{n-1}e^{-uk}\ln\left(\frac{ae^{-u+\frac{u}{a}}}{b}\right)du$$

$$= \ln\frac{\Gamma m}{k} - (m-1)\Psi(n) + \frac{n}{k}(k-1) - \frac{k^n}{\Gamma n}\ln(\frac{a}{b})\int_0^\infty u^{n-1}e^{-uk}du - \frac{k^n}{\Gamma n}\int_0^\infty u^{n-1}e^{-uk}(\frac{1}{a}-1)udu.$$

This gives

$$H(f_{n,k}, f_{m,k}) = \ln\frac{\Gamma m}{k} - (m-1)\Psi(n) + n - \ln(\frac{a}{b}) - \frac{n}{ak}. \tag{4.3.8}$$

When $m=k=1$, (4.3.8) reduces to

$$H(f_n, f) = \ln(\Gamma m) - (m-1)\Psi(n) + n - \ln(\frac{a}{b}) - \frac{n}{a}.$$

the inaccuracy measure between n^{th} and m^{th} record value for finite range distribution.

When $n = m$ and $k = 1$ (4.3.8) reduces to

$$H(f_n) = log(\Gamma n) - (n-1)\Psi(n) + n - log(\frac{a}{b}) - \frac{n}{a},$$

the shannon entropy of n^{th} record value for finite range distribution.
When $m = n = k = 1$, (4.3.8) becomes

$$H(f) = 1 - \ln(\frac{a}{b}) - \frac{1}{a},$$

the shannon entropy of parent distribution for finite range distribution.

4.4 Characterization Problem

For proving characterization result, we will use the lemma by Goffman and Pedrick [48] which we have already stated and used refer to Lemma 3.2.1. The result is as follows:

Theorem 4.4.1. *Let X and Y be two non-negative random variables having distribution function F and G. Let $H(f_{n,k}, f_{m,k}) < \infty$ and $H(g_{n,k}, g_{m,k}) < \infty$ are the inaccuracy measures between n^{th} and m^{th} upper k-record values for the parent distribution F and G respectively. Then F and G belong to the same location family of distribution, if and only if*

$$H(f_{n,k}, f_{m,k}) = H(g_{n,k}, g_{m,k}) \ , \ \forall\, n,k \geq 1.$$

Proof. The necessary part is clear. We need to prove the sufficiency part only. Let

$$H(f_{n,k}, f_{m,k}) = H(g_{n,k}, g_{m,k}) \ , \ \forall\, n,k \geq 1.$$

From (4.3.1), we know

$$H(f_{n,k}, f_{m,k}) = -\ln\frac{k^m}{\Gamma m} - (m-1)\Psi(n) + (m-1)\ln k + \frac{n}{k}(k-1) - \frac{k^n}{\Gamma n}.$$
$$\int_0^{\infty} u^{n-1} e^{-uk} \ln f(F^{-1}(1-e^{-u}))du,$$

where $u = -\ln \bar{F}(x)$ and hence $x = F^{-1}(1-e^{-u})$. Similarly we get

$$H(g_{n,k}, g_{m,k}) = -\ln\frac{k^m}{\Gamma m} - (m-1)\Psi(n) + (m-1)\ln k + \frac{n}{k}(k-1)$$
$$- \frac{k^n}{\Gamma n}\int_0^{\infty} u^{n-1} e^{-uk} \ln g(G^{-1}(1-e^{-u}))du.$$

Let $uk = z$, then the above equations become respectively

$$H(f_{n,k}, f_{m,k}) = -\ln\frac{k^m}{\Gamma m} - (m-1)\Psi(n) + (m-1)\ln k + \frac{n}{k}(k-1) \\ -\frac{1}{\Gamma n}\int_0^\infty z^{n-1}e^{-z}\ln\left(f(F^{-1}(1-e^{\frac{-z}{k}}))\right)dz \tag{4.4.1}$$

and

$$H(g_{n,k}, g_{m,k}) = -\ln\frac{k^m}{\Gamma m} - (m-1)\Psi(n) + (m-1)\ln k + \frac{n}{k}(k-1) \\ -\frac{1}{\Gamma n}\int_0^\infty z^{n-1}e^{-z}\ln\left(g(G^{-1}(1-e^{\frac{-z}{k}}))\right)dz. \tag{4.4.2}$$

Equating (4.4.1) and (4.4.2), we obtain

$$\frac{1}{\Gamma n}\int_0^\infty \left(\ln f(F^{-1}(1-e^{\frac{-z}{k}})) - \ln g(G^{-1}(1-e^{\frac{-z}{k}}))\right)e^{-z}z^{n-1}dz = 0,$$

It can be rewritten

$$\frac{1}{\Gamma n}\int_0^\infty \left(\ln f(F^{-1}(1-e^{\frac{-z}{k}})) - \ln g((G^{-1}(1-e^{\frac{-z}{k}}))\right)e^{-\frac{z}{2}}L_n(z)dz = 0, \tag{4.4.3}$$

for all $n \geq 1$, where $L_n(z)$ is the Laguerre polynomial defined in the Lemma 3.2.1. Using Lemma 3.2.1 and simplifying further, we obtain

$$f\left(F^{-1}(1-e^{-\frac{z}{k}})\right) = g\left(G^{-1}(1-e^{-\frac{z}{k}})\right), \ \forall n \geq 1.$$

Let $(1-e^{-\frac{z}{k}}) = v$. As $\frac{d}{dv}F^{-1}(v) = \frac{1}{f(F^{-1}(v))}$, it follows that

$$\begin{aligned} (F^{-1})^{'}(v) &= (G^{-1})^{'}(v) \quad \forall \ v \in (0,1). \\ \Rightarrow F^{-1}(v) &= G^{-1}(v) + a, \end{aligned}$$

where a is a constant. Hence the desired result follows. □

4.5 Residual Measure Of Inaccuracy For *k*-Records

Since the information measures defined previously are not applicable to a system which has already worked for some unit of time, thus the concept of residual measures has been developed in the literature. Let X denote lifetime of a system. The residual lifetime of the system when it is still working at time t, is denoted by $X_t = (X - t \mid X > t)$. Thus provided a system has survived up to time t, the corresponding dynamic measures of uncertainty, refer to Ebrahimi [36], and of discrimination, refer to Ebrahimi and Kirmani [40], are given as

$$H(f;t) = -\int_t^\infty \frac{f(x)}{\bar{F}(t)} \ln \frac{f(x)}{\bar{F}(t)} dx$$

and

$$H(f \mid g;t) = \int_t^\infty \frac{f(x)}{\bar{F}(t)} \ln \frac{f(x)/\bar{F}(t)}{g(x)/\bar{G}(t)} dx$$

respectively. Further Taneja et al. [117] introduced a dynamic measure of inaccuracy between two residual lifetime distributions as

$$H(f,g;t) = -\int_t^\infty \frac{f(x)}{\bar{F}(t)} \ln \frac{g(x)}{\bar{G}(t)} dx\,. \tag{4.5.1}$$

When $t = 0$, it reduces to Kerridge inaccuracy measure [63]. Madadi and Tata [77] gave the generalized results for Shannon information measure to k-records. Asha and Chacko [14] have studied residual Renyi entropy for k-record values. Now, corresponding to the measure (4.5.1), we propose the residual measure of inaccuracy between the n^{th} upper k-record value and the parent distribution as

$$H(f_{n,k}, f;t) = -\int_t^\infty \frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)} \ln \frac{f(x)}{\bar{F}(t)} dx\,, t > 0. \tag{4.5.2}$$

Next, we give a characterization result for the above measure and for proving it we will first state the following theorem and lemma from Gupta and Kirmani [51] which we have also used earlier.

Theorem 4.5.1. *Let the function f be a continuous function defined in a domain $D \subset R^2$ and f satisfies the Lipschitz condition (with respect to y) in D, that is*

$$|f(x,y_1) - f(x,y_2| \le k|y_1 - y_2|, k > 0,$$

for every point (x,y_1) and (x,y_2) in D. Then the function $y = \phi(x)$ satisfying the initial value problem $y' = f(x,y)$ and $\phi(x_0) = y_0$, $x \varepsilon I$, is unique.

Lemma 4.5.1. *Suppose that the function f is continuous in a convex region $D \subset R^2$. Suppose that $\frac{\partial f}{\partial y}$ exists and it is continuous in D. Then, f satisfies Lipschitz condition in D.*

Theorem 4.5.2. *Let X be a non-negative continuous random variable with distribution function F(x). Let the residual measure of inaccuracy of the corresponding n^{th} k-record* *value , denoted by $H(f_{n,k},f;t)$ be finite for all $t \geq 0$. Then $H(f_{n,k},f;t)$ characterizes the distribution.*

Proof. We know that

$$\begin{aligned} H(f_{n,k},f;t) &= -\int_t^{\infty} \frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)} \ln \frac{f(x)}{\bar{F}(t)} dx\,, \quad t > 0 \\ &= \ln \bar{F}(t) - \frac{1}{\bar{F}_{n,k}(t)} \int_t^{\infty} f_{n,k}(x) \ln f(x) dx. \end{aligned}$$

Taking derivative of both sides with respect to t, we have

$$\frac{d}{dt}(H(f_{n,k},f;t)) = -\lambda_F(t) + \lambda_{F_{n,k}}(t)(H(f_{n,k},f;t)) + \ln(\lambda_F(t)),$$

where $\lambda_F(t)$ and $\lambda_{F_{n,k}}(t)$ are the hazard rates of X and $X_{n,k}$ respectively. Taking derivative with respect to t again and using the relation

$$\lambda_{F_{n,k}}(t) = \alpha(t)\lambda_F(t).$$

where

$$\alpha(t) = \left(\frac{k^n(-\ln \bar{F}(t))^{n-1}(\bar{F}(t))^k}{\Gamma n, -k \ln \bar{F}(t)} \right),$$

we get

$$\lambda_F'(t) = \frac{\alpha(t)\lambda_F(t)H'(f_{n,k},f;t) + \alpha'(t)\lambda_F(t)H(f_{n,k},f;t) + \alpha'(t)\lambda_F(t)\ln\lambda_F(t) - H"(f_{n,k},f;t)}{1-\alpha(t)-\alpha(t)H(f_{n,k},f;t)-\alpha(t)\ln\lambda_F(t)}. \tag{4.5.3}$$

Suppose there are two distribution functions F and F* such that

$$H(f_{n,k},f;t) = H(f^*_{n,k},f;t) = c(t), say.$$

Then for all t, from (4.5.3) we get

$$\lambda_F'(t) = \psi(t,\lambda_F(t)), \lambda_{F*}'(t) = \psi(t,\lambda_{F*}(t)),$$

where

$$\psi(t,y) = \left(\frac{\alpha(t)c'(t)y + \alpha'(t)c(t)y + \alpha'(t)y\ln y - c"(t)}{1-\alpha(t)-\alpha(t)c(t)-\alpha(t)\ln y} \right).$$

Using Theorem 4.5.1 and Lemma 4.5.1 we get, $\lambda_F(t) = \lambda_{F*}(t)$, for all t. Also hazard rate function characterizes the distribution function uniquely, so we get the desired result. □

Lemma 4.5.2. *Let $\{X_i, i \geq 1\}$ be a sequence of i.i.d. continuous random variables having distribution F(x) and pdf f(x). Let $X_{n,k}$ be the n^{th} upper k-record value. Then the residual measure of inaccuracy (4.5.1) between n^{th} upper k-record and parent distribution can be represented as*

$$H(f_{n,k}, f;t) = \ln \bar{F}(t) - \frac{1}{\bar{F}_{n,k}(t)\Gamma n} \int_{-k\ln \bar{F}(t)}^{\infty} u^{n-1} e^{-u} \ln f(F^{-1}(1-e^{-\frac{u}{k}}))du. \qquad (4.5.4)$$

Proof. Consider

$$H(f_{n,k}, f;t) = -\int_t^{\infty} \frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)} \ln \frac{f(x)}{\bar{F}(t)} dx \;, t > 0$$

$$= -\frac{1}{\bar{F}_{n,k}(t)} \int_t^{\infty} f_{n,k}(x) \ln f(x) dx + \ln \bar{F}(t).$$

Using (1.10.1)

$$H(f_{n,k}, f;t) = -\frac{1}{\bar{F}_{n,k}(t)} \int_t^{\infty} \frac{k^n}{\Gamma n} [-\ln \bar{F}(x)]^{n-1} [\bar{F}(x)]^{k-1} f(x) \ln f(x) dx + \ln \bar{F}(t).$$

and putting $-k \ln \bar{F}(x) = u$, we get

$$H(f_{n,k}, f;t) = \ln \bar{F}(t) - \frac{1}{\bar{F}_{n,k}(t)\Gamma n} \int_{-k\ln \bar{F}(t)}^{\infty} u^{n-1} e^{-u} \ln f(F^{-1}(1-e^{-\frac{u}{k}}))du.$$

□

Example 4.5.1. *The pdf of the finite range distribution is given by*

$$f(x) = \frac{a}{b}\left(1-\frac{x}{b}\right)^{a-1}, \quad a > 1, \;\; 0 \leq x \leq \; b.$$

The survival function is

$$\bar{F}(x) = 1 - F(x) = \left(1-\frac{x}{b}\right)^{a}.$$

Substituting $-k \ln \bar{F}(x) = u$, we observe that $x = F^{-1}(1-e^{-\frac{u}{k}}) = b(1-e^{-\frac{u}{ak}})$.

and

$$f\left(F^{-1}(1-e^{-\frac{u}{k}})\right) = f\left(b(1-e^{-\frac{u}{ak}})\right) = \frac{a}{b}\, e^{-\frac{u}{k}(1-\frac{1}{a})}.$$

Putting in (4.5.4), we will get

$$H(f_{n,k}, f;t) = a\ln(1-\frac{t}{b}) - \ln(\frac{a}{b}) + \frac{\Gamma(n+1, -k\ln F\bar(t))}{\Gamma(n, -k\ln F\bar(t))} \left(\frac{a-1}{ak}\right).$$

Proposition 4.5.1. *If $H(f_{n,k}, f)$ is the inaccuracy measure between n^{th} k-record value* *and parent random variable, then*

$$H(f_{n,k}, f; t) \leq \frac{H(f_{n,k}, f)}{\bar{F}_{n,k}(t)}.$$

Proof. We know

$$H(f_{n,k}, f; t) = \ln \bar{F}(t) - \frac{1}{\bar{F}_{n,k}(t)} \int_t^\infty f_{n,k}(x) \ln f(x) dx \ , \ t > 0$$

As for $t \geq 0 \quad \ln \bar{F}(t) \leq 0$

$$\begin{aligned}
\Rightarrow H(f_{n,k}, f; t) &\leq -\frac{1}{\bar{F}_{n,k}(t)} \int_t^\infty f_{n,k}(x) \ln f(x) dx \ , \ t > 0 \\
&\leq -\frac{1}{\bar{F}_{n,k}(t)} \int_0^\infty f_{n,k}(x) \ln f(x) dx \ , \\
&= \frac{H(f_{n,k}, f)}{\bar{F}_{n,k}(t)}.
\end{aligned}$$

□

Proposition 4.5.2. *Let M be the mode of the distribution.Then*

$$H(f_{n,k}, f; t) \geq \ln \frac{\bar{F}(t)}{M}.$$

Proof. As M is the mode, therefore $f(x) \leq M, \ \forall \, x \geq t$.Therefore,

$$\begin{aligned}
H(f_{n,k}, f; t) &= \ln \bar{F}(t) - \frac{1}{\bar{F}_{n,k}(t)} \int_t^\infty f_{n,k}(x) \ln f(x) dx \ , \ t > 0 \\
&\geq \ln \bar{F}(t) - \frac{\ln M}{\bar{F}_{n,k}(t)} \int_t^\infty f_{n,k}(x) dx \ , \ t > 0 \\
&= \ln \frac{\bar{F}(t)}{M}.
\end{aligned}$$

This proves the result. □

4.6 Conclusion

The inaccuracy measure has various applications in different areas of science and technology such as statistical inference, estimation and coding theory. We have studied the inaccuracy measure based on k-record values. The inaccuracy measure for various distributions which are commonly used in the reliability modeling, has been discussed. A characterization result has also been given. We have also studied the concept of residual measure of inaccuracy for k-record values and also some properties of it. It has also shown that it characterizes the distribution function uniquely.

Chapter 5

Cumulative Residual Inaccuracy Measure For *k*-Record Values

5.1 Introduction

So far we have discussed information measure based on probability density function which have their own limitations. To overcome some of the limitations of Shannon's entropy measure as discussed in Section 1.7, a new measure of uncertainty has been developed by Rao et al. [95], known as cumulative residual entropy (CRE) in which the probability density function $f(x)$ has been replaced with the survival function $\bar{F}(x)$ of the random variable X, given as

$$H(\bar{F}) = -\int_0^\infty \bar{F}(x)\log\bar{F}(x)dx. \tag{5.1.1}$$

This measure of uncertainty is particularly appropriate to describe the information in problems connected with the ageing properties of reliability theory based on the mean residual life function. After Rao [95] proposed this cumulative residual entropy, this measure became the subject of interest for various researchers. Asadi and Zohrevand [13] considered its dynamic version to explain the age effect

The result of this chapter has been communicated in a research paper under the title **Cumulative Residual Inaccuracy Measure for *k*-Record Values** and some work has been presented in International Conference on Recent Advances in Pure and Applied Mathematics held at Delhi Technological University, Delhi, Oct 23-25, 2018.

on the information concerning the residual lifetime of a system or a component . In analogy with cumulative residual entropy (5.1.1), Dicrescenzo and Longobardi [34] introduced the cumulative entropy particularly suitable for the problems related to inactivity time. Sunoj and Linu [109] proposed cumulative Renyi's entropy. Taneja and Kumar [116] extended the concept of cumulative Residual Entropy to cumulative residual inaccuracy and then to dynamic cumulative inaccuracy and also studied some of its properties. The measure proposed by Taneja and Kumar [116] is given as

$$H(\bar{F},\bar{G}) = -\int_0^\infty \bar{F}(x)\log\bar{G}(x)dx. \tag{5.1.2}$$

Psarrakos and Navarro [93] studied this concept for k-record values. Tahmasebi and Eskandarzadeh [112] proposed the extension of cumulative entropy based on k^{th} lower record values and also considered its dynamic version using the past lifetime.

In the preceding chapters we have studied the inaccuracy measure between record distribution and the parent distribution and then between k-record distribution and parent distribution. In this chapter we study the cumulative residual inaccuracy contained in the sequence of k-record values. The organisation of the chapter is as follows: In Section 5.2, we propose the extension of cumulative residual inaccuracy measure to k-record values. In Section 5.3, we study some of the properties of the proposed measure and find some bounds to the measure. Then in Section 5.4, some stochastic ordering has been studied. Section 5.5 provides the simplified expression for the cumulative residual inaccuracy to make the computations and calculations easy. Section 5.6 concludes the chapter.

5.2 Cumulative Residual Inaccuracy Measure

Corresponding to the measure (5.1.2), here we are introducing the cumulative residual inaccuracy measure between k-record distribution $\bar{F}_{n,k}(x)$ and the parent distribution $\bar{F}$ as

$$H(\bar{F}_{n,k},\bar{F}) = -\int_0^\infty \bar{F}_{n,k}(x)\ln\bar{F}(x)dx. \tag{5.2.1}$$

Using (1.10.3)

$$H(\bar{F}_{n,k},\bar{F}) = -\int_0^\infty \sum_{j=0}^{n-1} \frac{1}{j!} (\bar{F}(x))^k (-k\ln \bar{F}(x))^j \ln \bar{F}(x) dx$$
$$= \sum_{j=0}^{n-1} \int_0^\infty \frac{k^j}{j!} (\bar{F}(x))^k (-\ln \bar{F}(x))^{j+1} dx. \tag{5.2.2}$$

After some rearrangements, we get

$$\begin{aligned} H(\bar{F}_{n,k},\bar{F}) &= \sum_{j=0}^{n-1} \int_0^\infty \frac{(j+1)}{k^2 \lambda_F(x)} . \frac{k^{j+2} (-\ln \bar{F}(x))^{(j+1)} f(x)}{\Gamma(j+2)} dx \\ &= \sum_{j=0}^{n-1} \int_0^\infty \frac{(j+1)}{k^2 \lambda_F(x)} f_{j+2,k}(x)\, dx \\ &= \sum_{j=0}^{n-1} \frac{(j+1)}{k^2} E_{f_{j+2,k}} \left(\frac{1}{\lambda_F(x)} \right). \end{aligned} \tag{5.2.3}$$

Here E stands for expectation and $\lambda_F(x)$ stands for hazard rate function corresponding to $F(x)$.

5.3 Properties And The Bounds To The Measure

In the present section, we study some of the properties of the proposed measure of cumulative inaccuracy as follows:

1. If $\mu_{n,k}(x) = \int_0^\infty \bar{F}_{n,k} dx$, then inaccuracy measure can be expressed as

$$H(\bar{F}_{n,k},\bar{F}) = \sum_{j=0}^{(n-1)} k^j (j+1) \left(\mu_{j+2,k}(x) - \mu_{j+1,k}(x) \right),$$

Proof. From (1.10.3), we can write

$$\bar{F}_{j+2,k}(x) - \bar{F}_{j+1,k}(x) = (\bar{F}(x))^k \frac{(-k\ln \bar{F}(x))^{j+1}}{(j+1)!}. \tag{5.3.1}$$

Therefore from (5.2.2) and (5.3.1), we get

$$\begin{aligned} H(\bar{F}_{n,k},\bar{F}) &= \sum_{j=0}^{(n-1)} k^j (j+1) \int_0^\infty \left(F_{j+2,k}(x) - F_{j+1,k}(x) \right) dx \\ &= \sum_{j=0}^{(n-1)} k^j (j+1) \left(\mu_{j+2,k}(x) - \mu_{j+1,k}(x) \right). \end{aligned}$$

Remark 5.3.1. *If for a fixed* k, $F_{n,k}$ *is a decreasing function of* n, *that is,* $\bar{F}_{n,k}$ *is an increasing function of* n, *then* $\bar{F}_{j+2,k} > \bar{F}_{j+1,k}$. *From above result, we can see that* $H(\bar{F}_{n,k}, \bar{F})$ *is an increasing function of* n.

2. Consider two random variables X and Y with survival functions $\bar{F}(x)$ and $\bar{G}(y)$ respectively such that $Y = \phi(X)$, where ϕ is a strictly increasing function with $\phi(0) = 0$, then

$$H(\bar{G}_{n,k}, \bar{G}) = \sum_{j=0}^{(n-1)} \int_0^\infty \frac{k^j}{j!} (\bar{F}(x))^k (-\ln \bar{F}(x))^{j+1} \phi^{'}(x) dx. \tag{5.3.2}$$

Proof. We can write from (5.2.2)

$$H(\bar{G}_{n,k}, \bar{G}) = \sum_{j=0}^{(n-1)} \int_0^\infty \frac{k^j}{j!} (\bar{G}(y))^k \left(-\ln \bar{G}(y)\right)^{j+1} dy. \tag{5.3.3}$$

Now $Y = \phi(X) \Rightarrow \bar{G}(y) = \bar{F}(x)$ and $\bar{G}_{n,k}(y) = \bar{F}_{n,k}(x)$. Also $dy = \phi^{'}(x)dx$. Putting all these values in (5.3.3), the result is obvious. □

Remark 5.3.2. *In particular* $Y = \phi(X) = aX \Rightarrow \phi^{'}(x) = a$. *Therefore (5.3.2) becomes*

$$\begin{aligned} H(\bar{G}_{n,k}, \bar{G}) &= \sum_{j=0}^{(n-1)} a \int_0^\infty \frac{k^j}{j!} (\bar{F}(x))^k (-\ln \bar{F}(x))^{j+1} dx \\ &= a\, H(\bar{F}_{n,k}, \bar{F}). \end{aligned}$$

3. If $\bar{G}(x) = (\bar{F}(x))^\beta$, β is an integer greater than 1, then

$$H(\bar{G}_{n,k}, \bar{G}) = \beta H(\bar{F}_{n,k\beta}, \bar{F})$$

Proof. We know

$$\begin{aligned} H(\bar{G}_{n,k}, \bar{G}) &= \sum_{j=0}^{(n-1)} \int_0^\infty \frac{k^j}{j!} (\bar{G}(x))^k (-\ln \bar{G}(x))^{j+1} dx \\ &= \sum_{j=0}^{(n-1)} \int_0^\infty \frac{k^j}{j!} (\bar{F}(x))^{k\beta} (-\beta \ln \bar{F}(x))^{j+1} dx \\ &= \beta \sum_{j=0}^{(n-1)} \int_0^\infty \frac{(k\beta)^j}{j!} (\bar{F}(x))^{k\beta} (-\ln \bar{F}(x))^{j+1} dx \\ &= \beta H(\bar{F}_{n,k\beta}, \bar{F}). \end{aligned}$$

4. Consider $\eta(X) = -\int_0^\infty (\bar{F}(x))^k \ln \bar{F}(x) dx$, then

$$H(\bar{F}_{n,k}, \bar{F}) \geq \sum_{j=0}^{(n-1)} \frac{k^j}{j!} (\eta(X))^{j+1}, \tag{5.3.4}$$

Proof. From (5.2.2)

$$\begin{aligned} H(\bar{F}_{n,k}, \bar{F}) &= \sum_{j=0}^{(n-1)} \int_0^\infty \frac{k^j}{j!} (\bar{F}(x))^k (-\ln \bar{F}(x))^{j+1} dx \\ &= \sum_{j=0}^{(n-1)} \int_0^\infty \frac{k^j}{j!} (\bar{F}(x))^{k-1} \bar{F}(x) (-\ln \bar{F}(x))^{j+1} dx \\ &\geq \sum_{j=0}^{(n-1)} \int_0^\infty \frac{k^j}{j!} (\bar{F}(x))^{k-1} (\bar{F}(x))^{j+1} (-\ln \bar{F}(x))^{j+1} dx. \end{aligned}$$

Now, using Jensen's inequality, we get

$$\begin{aligned} H(\bar{F}_{n,k}, \bar{F}) &\geq \sum_{j=0}^{(n-1)} \frac{k^j}{j!} \left(\int_0^\infty - (\bar{F}(x))^k \ln \bar{F}(x) \right)^{j+1} dx \\ &\geq \sum_{j=0}^{(n-1)} \frac{k^j}{j!} (\eta(X))^{j+1}. \end{aligned}$$

Here $\eta(X) = -\int_0^\infty (\bar{F}(x))^k \ln \bar{F}(x) dx$. □

5. Let X be an absolutely continuous non-negative random variable, then

$$H(\bar{F}_{n,1}, \bar{F}) \geq \sum_{j=0}^{(n-1)} \frac{(H(\bar{F}))^{j+1}}{j!}, \tag{5.3.5}$$

where $H(\bar{F})$ is given by (5.1.1).

Proof. This result can be proved directly from (5.3.4) by taking $k = 1$. Also if we put $k = 1$ in $\eta(X) = -\int_0^\infty (\bar{F}(x))^k \ln \bar{F}(x) dx$, it becomes the cumulative residual entropy given as $H(\bar{F}) = -\int_0^\infty \bar{F}(x) \ln \bar{F}(x) dx$. This proves the result. □

5.4 Some Results On Stochastic Ordering

In this section we prove some order properties of cumulative inaccuracy measure for k-record values. First we give following definitions.

Definition 5.4.1. *A random variable X is said to be less than Y in the stochastic ordering denoted by $X \overset{st}{\leq} Y$ if $\bar{F}(x) \leq \bar{G}(x)$ for all x, where $\bar{F}(x)$ and $\bar{G}(x)$ are the survival functions of X and Y respectively.*

Definition 5.4.2. *A random variable X is said to be less than Y in the likelihood ratio ordering denoted by $X \overset{lr}{\leq} Y$ if $\frac{f_X(x)}{g_Y(x)}$ is non increasing in x, where $f_X(x)$ and $g_Y(x)$ are the pdf of X and Y respectively.*

Proposition 5.4.1. *If $E(X_{n,k})$ and $E(X)$, are the expected value of n^{th} k-record value and the parent distribution such that $X_{n,k} \overset{st}{\leq} X$, then*

$$(i)\ H(\bar{F}_{n,k}) \leq H(\bar{F}_{n,k}, \bar{F}) - E(X_{n,k}) \ln \frac{E(X_{n,k})}{E(X)}. \tag{5.4.1}$$

$$(ii)\ H(\bar{F}_{n,k}) \leq H(\bar{F}) - E(X_{n,k}) \ln \frac{E(X_{n,k})}{E(X)}. \tag{5.4.2}$$

Here $H(\bar{F}_{n,k})$ and $H(\bar{F})$ denote the cumulative residual entropy for the random variables $X_{n,k}$ and X respectively.

Proof. Using log-sum inequality we can write

$$\begin{aligned}\int_0^\infty \bar{F}_{n,k}(x) \ln \frac{\bar{F}_{n,k}(x)}{\bar{F}(x)} dx &\geq \int_0^\infty \bar{F}_{n,k}(x) dx\ \ln \frac{\int_0^\infty \bar{F}_{n,k}(x) dx}{\int_0^\infty \bar{F}(x) dx} \\ &= E(X_{n,k}) \ln \frac{E(X_{n,k})}{E(X)}.\end{aligned}$$

Hence using above inequality, we obtain

$$\begin{aligned}H(\bar{F}_{n,k}) &= -\int_0^\infty \bar{F}_{n,k}(x) \ln \bar{F}_{n,k}(x)\, dx \\ &\leq -\int_0^\infty \bar{F}_{n,k}(x) \ln \bar{F}(x)\, dx - E(X_{n,k}) \ln \frac{E(X_{n,k})}{E(X)} \\ &= H(\bar{F}_{n,k}, \bar{F}) - E(X_{n,k}) \ln \frac{E(X_{n,k})}{E(X)}.\end{aligned}$$

Now using $X_{n,k} \overset{st}{\leq} X$ in above inequality, we get

$$\begin{aligned}H(\bar{F}_{n,k}) &\leq -\int_0^\infty \bar{F}_{n,k}(x) \ln \bar{F}(x)\, dx - E(X_{n,k}) \ln \frac{E(X_{n,k})}{E(X)} \\ &\leq -\int_0^\infty \bar{F}(x) \ln \bar{F}(x)\, dx - E(X_{n,k}) \ln \frac{E(X_{n,k})}{E(X)} \\ &= H(\bar{F}) - E(X_{n,k}) \ln \frac{E(X_{n,k})}{E(X)}.\end{aligned}$$

This proves the result. □

Proposition 5.4.2. *Let $X > 0$ be with density function $f(x)$ and cumulative distribution function $F(x)$. If $X_{n,k} \overset{st}{\leq} X$, then*

$$H(\bar{F}_{n,k}, \bar{F}) \leq Ce^{H(f_{n,k}, f)}. \tag{5.4.3}$$

Here $H(f_{n,k}, f) = -\int_0^\infty f_{n,k}(x) \ln f(x) dx$.

Proof. Consider

$$\begin{aligned}\int_0^\infty f_{n,k}(x) \ln \frac{f(x)}{\bar{F}_{n,k}(x) \ln \bar{F}(x)} dx &\geq \ln \frac{1}{\int_0^\infty \bar{F}_{n,k}(x) \ln \bar{F}(x) dx} \\ &= \ln \frac{1}{-H(\bar{F}_{n,k}, \bar{F})}.\end{aligned}$$

The inequality above results from log-sum inequality. Continuing, we get

$$\int_0^\infty f_{n,k}(x) \ln f(x) dx - \int_0^\infty f_{n,k}(x) \ln \left(\bar{F}_{n,k}(x) \ln \bar{F}(x) \right) dx \geq - \ln \left(-H(\bar{F}_{n,k}, \bar{F}) \right).$$

or

$$H(f_{n,k}, f) + \int_0^\infty f_{n,k}(x) \ln \left(\bar{F}_{n,k}(x) \ln \bar{F}(x) \right) dx \leq \ln(-H(\bar{F}_{n,k}, \bar{F})).$$

Using $X_{n,k} \overset{st}{\leq} X$, we get

$$H(f_{n,k}, f) + \int_0^\infty f_{n,k}(x) \ln \left(\bar{F}_{n,k}(x) \ln \bar{F}_{n,k}(x) \right) dx \leq \ln(-H(\bar{F}_{n,k}, \bar{F})). \tag{5.4.4}$$

Using substitution $\bar{F}_{n,k}(x) = u$, we get

$$\int_0^\infty f_{n,k}(x) \ln \left(\bar{F}_{n,k}(x) \ln \bar{F}_{n,k}(x) \right) dx = \int_0^1 \ln (u \ln u) du = k. \ (say)$$

Therefore, by putting this value in (5.4.4)

$$H(f_{n,k}, f) + k \leq \ln(-H(\bar{F}_{n,k}, \bar{F})).$$

or

$$H(\bar{F}_{n,k}, \bar{F}) \leq C\, e^{H(f_{n,k}, f)},$$

where C denotes the constant. □

Proposition 5.4.3. *Let X be a non-negative random variable, then*

$$H(\bar{F}_{n,k}) \leq \frac{E_{f_{n,k}}(X^{\beta+1})\Gamma^{\beta}(1+\frac{1}{\beta})}{(\beta+1)E^{\beta}(X_{n,k})}. \tag{5.4.5}$$

Proof. Let X has Weibull distribution with reliability function $\bar{F}(x) = e^{-(\lambda x)^{\beta}}$.
From (5.4.1)

$$H(\bar{F}_{n,k}) \leq -\int_0^{\infty} \bar{F}_{n,k}(x) \ln \bar{F}(x)\, dx - E(X_{n,k}) \ln \frac{E(X_{n,k})}{E(X)}.$$

For Weibull distribution, it becomes

$$\begin{aligned} -H(\bar{F}_{n,k}) &\geq E(X_{n,k}) \ln \frac{E(X_{n,k})}{E(X)} - \int_0^{\infty} \bar{F}_{n,k}(x)(\lambda x)^{\beta} dx \\ &= E(X_{n,k}) \ln \frac{E(X_{n,k})}{E(X)} - \frac{\lambda^{\beta}}{\beta+1} E_{f_{n,k}}(X^{\beta+1}). \end{aligned}$$

Let $E(X) = \mu = \int_0^{\infty} \bar{F}(x)dx = \frac{\Gamma(\frac{1}{\beta}+1)}{\lambda}$. *Hence,*

$$-H(\bar{F}_{n,k}) \geq E(X_{n,k}) \ln \frac{E(X_{n,k})}{\mu} - \frac{E_{f_{n,k}}(X^{\beta+1})\Gamma^{\beta}(1+\frac{1}{\beta})}{(\beta+1)\mu^{\beta}}. \tag{5.4.6}$$

The right hand side of above equation is maximized for a fixed β *at*

$$\mu_{\beta} = \left(\frac{\beta E_{f_{n,k}}(X^{\beta+1})\Gamma^{\beta}(1+\frac{1}{\beta})}{(\beta+1)E(X_{n,k})} \right)^{\frac{1}{\beta}}.$$

Using this in (5.4.6)

$$\begin{aligned} -H(\bar{F}_{n,k}) &\geq E(X_{n,k}) \ln \frac{E(X_{n,k})}{\mu_{\beta}} - \frac{E_{f_{n,k}}(X^{\beta+1})\Gamma^{\beta}(1+\frac{1}{\beta})}{(\beta+1)\mu_{\beta}^{\beta}} \\ &= -\frac{E(X_{n,k})}{\beta} \ln \left(\frac{\beta E_{f_{n,k}}(X^{\beta+1})\Gamma^{\beta}(1+\frac{1}{\beta})}{(\beta+1)E^{\beta+1}(X_{n,k})} \right) - \frac{E(X_{n,k})}{\beta}. \\ &\geq \frac{E(X_{n,k})}{\beta} \left(1 - \frac{\beta E_{f_{n,k}}(X^{\beta+1})\Gamma^{\beta}(1+\frac{1}{\beta})}{(\beta+1)E^{\beta+1}(X_{n,k})} \right) - \frac{E(X_{n,k})}{\beta} \\ &= \frac{-E_{f_{n,k}}(X^{\beta+1})\Gamma^{\beta}(1+\frac{1}{\beta})}{(\beta+1)E^{\beta}(X_{n,k})}. \\ \text{or} \quad H(\bar{F}_{n,k}) &\leq \frac{E_{f_{n,k}}(X^{\beta+1})\Gamma^{\beta}(1+\frac{1}{\beta})}{(\beta+1)E^{\beta}(X_{n,k})}. \end{aligned}$$

This completes the proof.

Remark 5.4.1. *If we put $\beta = 1$ and $n = k = 1$ in above result, it reduces to*

$$H(\bar{F}) \leq \frac{E(X^2)}{2E(X)}, \tag{5.4.7}$$

a bound obtained by Rao et al. [95].

Proposition 5.4.4. *Suppose that the non-negative random variable X has decreasing hazard rate, then*

$$H(\bar{F}_{n+1,k}, \bar{F}) \leq H(\bar{F}_{n,k}, \bar{F}) \tag{5.4.8}$$

Proof. Consider two pdfs of consecutive record values $f_{n,k}(x)$ and $f_{n+1,k}(x)$.
Using (1.10.1), we get

$$\frac{f_{n,k}(x)}{f_{n+1,k}(x)} = -\frac{n}{k \ln \bar{F}(x)}, \tag{5.4.9}$$

which is a decreasing function in x. This implies that $X_{n,k} \overset{lr}{\leq} X_{n+1,k}$. Therefore $X_{n,k} \overset{st}{\leq} X_{n+1,k}$, that is $\bar{F}_{n,k}(x) \leq \bar{F}_{n+1,k}(x)$. (For more details one can refer to Shaked and Shantikumar [103]). Therefore for all increasing function ψ, this is equivalent to $E(\psi(X_{n,k})) \leq E(\psi(X_{n+1,k}))$, provided these expectations exist.
Now if X has decreasing hazard rate $\lambda_F(x)$, then $\frac{1}{\lambda_F(x)}$ is an increasing function.Therefore by above

$$E\left(\frac{1}{\lambda_F(X_{n,k})}\right) \leq E\left(\frac{1}{\lambda_F(X_{n+1,k})}\right)$$

From (5.2.3), we can see that

$$H(\bar{F}_{n,k}, \bar{F}) \leq H(\bar{F}_{n+1,k}, \bar{F}) \tag{5.4.10}$$

Here $\lambda_F(X_{n,k})$ and $\lambda_F(X_{n+1,k})$ denotes the hazard rate corresponding to $X_{n,k}$ and $X_{n+1,k}$ respectively. This completes the proof. □

5.5 Cumulative Inaccuracy For Some Specific Distributions

In this section first we give a lemma providing a simplified expression for finding the cumulative inaccuracy measure for various distributions and then we give some examples based on it.

Lemma 5.5.1. *Consider a random variable X having distribution function $\bar{F}(x)$, then cumulative inaccuracy measure between k-record values and parent distribution is given as*

$$H(\bar{F}_{n,k},\bar{F}) = \sum_{j=0}^{(n-1)} \frac{k^j}{j!}\int_0^\infty \frac{u^{j+1}e^{-u(k+1)}}{f(F^{-1}(1-e^{-u}))}du. \tag{5.5.1}$$

Proof. From (5.2.2)

$$H(\bar{F}_{n,k},\bar{F}) = (\bar{F}(x))^k \sum_{j=0}^{(n-1)} \int_0^\infty \frac{k^j}{j!}(-\ln \bar{F}(x))^{j+1}dx.$$

By putting $-\ln \bar{F}(x) = u$ in above equation, we get

$$\begin{aligned} H(\bar{F}_{n,k},\bar{F}) &= e^{-ku}\sum_{j=0}^{(n-1)} \int_0^\infty \frac{k^j u^{j+1}e^{-u}}{j!f(F^{-1}(1-e^{-u}))}du \\ &= \sum_{j=0}^{(n-1)} \frac{k^j}{j!}\int_0^\infty \frac{u^{j+1}e^{-u(k+1)}}{f(F^{-1}(1-e^{-u}))}du. \end{aligned}$$

□

Example 5.5.1. *Consider finite range distribution with pdf $f(x) = \frac{a}{b}\left(1-\frac{x}{b}\right)^{a-1}$, $a > 1$, $0 \le x \le b$ and distribution function $F(x) = \left(1-\frac{x}{b}\right)^a$.*
Then $F^{-1}(1-e^{-u}) = b(1-e^{-\frac{u}{a}})$ and this gives $f(F^{-1}(1-e^{-u})) = \frac{a}{b}e^{-\frac{u(a-1)}{a}}$. Putting all these values in (5.5.1), we get

$$\begin{aligned} H(\bar{F}_{n,k},\bar{F}) &= \sum_{j=0}^{(n-1)} \frac{k^j}{j!}\int_0^\infty \frac{u^{j+1}e^{-u(k+1)}}{f(F^{-1}(1-e^{-u}))}du \\ &= \sum_{j=0}^{(n-1)} \frac{k^j b}{j!a}\int_0^\infty \frac{u^{j+1}e^{-u(k+1)}}{e^{-\frac{u(a-1)}{a}}}du \\ &= \sum_{j=0}^{(n-1)} \frac{k^j b}{j!a}\int_0^\infty u^{j+1}e^{-u(k+\frac{1}{a})}du. \end{aligned}$$

Now using substitution $u(k+\frac{1}{a}) = t$

$$\begin{aligned} H(\bar{F}_{n,k},\bar{F}) &= \sum_{j=0}^{(n-1)} \frac{k^j b}{j!a(k+\frac{1}{a})^{j+2}}\int_0^\infty t^{j+1}e^{-t}dt \\ &= \frac{b}{a}\sum_{j=0}^{(n-1)} \frac{k^j (j+1)}{(k+\frac{1}{a})^{j+2}}. \end{aligned} \tag{5.5.2}$$

Here $\int_0^\infty t^{j+1}e^{-t}dt = \Gamma(j+2)$

In particular if $n=2$ *and* $k=1$*, Then we have inaccuracy measure between second record value and parent distribution as*

$$H(\bar{F}_{2,1},\bar{F}) = \frac{b}{a(1+\frac{1}{a})^2}\left(\sum_{j=1}^{3}\frac{j}{(1+\frac{1}{a})^{j-1}}\right). \tag{5.5.3}$$

Example 5.5.2. *For uniform distribution, if we put* $a=1$ *in (5.5.2), we get inaccuracy measure corresponding to uniform distribution as*

$$H(\bar{F}_{n,k},\bar{F}) = b\sum_{j=0}^{(n-1)}\frac{k^j(j+1)}{(k+1)^{j+2}}.$$

Example 5.5.3. *If* X *is a random variable with Weibull distribution having pdf* $f(x) = \alpha\beta x^{\beta-1}e^{-\alpha x^\beta}, \quad for\ x>0,\ \alpha>0,\ \beta>0$ *and survival function* $\bar{F}(x) = 1-e^{-\alpha x^\beta}$*, this gives* $F^{-1}(1-e^{-u}) = \left(\frac{u}{\alpha}\right)^{\frac{1}{\beta}}$*. Therefore putting these values in (5.5.1), then inaccuracy measure will come out to be*

$$H(\bar{F}_{n,k},\bar{F}) = \sum_{j=0}^{(n-1)}\frac{k^j\Gamma(j+1+\frac{1}{\beta})}{j!(\alpha\beta^\beta)^{\frac{1}{\beta}}k^{j+1+\frac{1}{\beta}}}. \tag{5.5.4}$$

Example 5.5.4. *If* X *is a exponentially distributed random variable, then by putting* $\beta=1$ *in (5.5.4), we get the inaccuracy measure corresponding to exponential distribution as*

$$H(\bar{F}_{n,k},\bar{F}) = \sum_{j=0}^{(n-1)}\frac{(j+1)}{k^2\alpha}. \tag{5.5.5}$$

Remark 5.5.1. *If we put* $k=1$*, then* $H(\bar{F}_{n,k},\bar{F})$ *becomes* $H(\bar{F}_n,\bar{F})$*, which represents the cumulative residual inaccuracy measure between* n^{th} *record value and parent distribution and if* $n=1,\ k=1$*, this represents the cumulative residual entropy given by (refer to Rao et al. [95]).*

5.6 Conclusion

Record values and k-record values originates very often in the many realistic situations. Athletic events, hydrology, weather forecasting are some of the examples. Also Kerridge [63] inaccuracy measure is quite useful to measure inaccuracy between two distributions. So by considering this, we have provided the cumulative residual inaccuracy between the distributions of k-record values and the parent random variable. Also we have obtained some properties of this measure including stochastic ordering. To minimise the computation work to get the inaccuracy for various distributions, we have provided a simplified expression for the proposed inaccuracy measure and have applied it to some of the standard distributions.

Chapter 6

Residual Measure Of Discrimination

6.1 Introduction

Sometimes in statistical analysis there is difference between registered distribution and the original distribution. Many events in nature are conveyed via the study of relation between registered and original distributions. Kullback Leibler [66] information measure is connected with the statistical problem of discrimination by considering a measure of distance or divergence between two distributions associated with the same random experiment. It is given by

$$H(f/g) = \int_0^\infty f(x) \ln \frac{f(x)}{g(x)} dx. \tag{6.1.1}$$

Here $f(x)$ is the actual distribution of random variable X and $g(x)$ is its predicted distribution. Basically this is a measure of how one probability distribution differs from a second probability distribution. If it comes out to be zero, it indicates that we can expect similar behaviour of two different distributions. Although above measure is not actual a metric, yet it is termed as 'distance' because $H(f/g) \geq 0$ and equality holds iff $f(x) = g(x), \ \forall\, x$. In various phenomenons for example life testing and survival analysis, one has the knowledge about the time 't' up to

The work presented in this chapter is communicated with the title **A Measure of Discrimination Between Two Residual Lifetime Distributions For k-Record Values** and some work has been presented in International Research Symposium on Engineering and Technology held at Singapore, August 28-30, 2018.

which the system has already worked. In such cases, this time 't' must be considered, while finding the discrimination measure between two systems. By taking in to account this fact, Ebrahimi and Kirmani [40] gave the measure of discrimination between the residual lifetime distributions by replacing the probability density functions $f(x)$ and $g(x)$ by corresponding residual pdfs $\frac{f(x)}{\bar{F}(t)}$ and $\frac{g(x)}{\bar{G}(t)}$ as

$$H(f/g;t) = \int_t^{\infty} \frac{f(x)}{\bar{F}(t)} \ln\left(\frac{f(x)/\bar{F}(t)}{g(x)/\bar{G}(t)}\right) dx. \tag{6.1.2}$$

Here $\bar{F}(t)$ and $\bar{G}(t)$ denote the survival functions corresponding to the distributions $f(x)$ and $g(x)$ respectively. When $t = 0,$ then the measure (6.1.2) becomes (6.1.1). Ebrahimi and Kirmani [39] studied a characterization result of measure of discrimination between two residual life distributions using proportional hazards model.

In the previous chapters we have discussed the inaccuracy measure between the distribution of n^{th} record value and the parent distribution and then the inaccuracy measure between n^{th} k-record value and the parent distribution. So it is natural that one can also extend the concept of K-L information measure and its dynamic version given by Ebrahimi and Kirmani [40] to k-record values. So considering this in the present chapter we provide an extension of the measure of discrimination given by Ebrahimi and kirmani [40] to k-record values. The distance between two k-record distributions of residual lifetime is found. Also keeping the record times fixed, we derive the distance between k-record value and l-record value.

The chapter is organised as follows : In Section 6.2, we propose a measure of discrimination between two k-record values of residual lifetime distribution and give a characterization result for that. In Section 6.3, we study some properties of this measure. In Section 6.4, the proposed measure is computed for some specific distributions. Some bounds to the discrimination measure are found in Section 6.5. Section 6.6 is of conclusion.

6.2 A Measure Of Discrimination Between Two Residual Lifetime Distributions

Record data appears in miscellaneous fields for example sports, hydrology, weather forecasting and even in the field of medicines. So the distance between two record value distributions becomes important. So taking this in to mind here we extend the measure of discrimination between two residual lifetime distributions proposed by Ebrahimi and Kirmani [40] to k- record values. For all $t>0$, we propose the measure of discrimination between n^{th} and m^{th} upper k-record values of residual lifetime as

$$\begin{aligned} H(f_{n,k}/f_{m,k};t) &= \int_t^\infty \frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)} \ln\left(\frac{\frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)}}{\frac{f_{m,k}(x)}{\bar{F}_{m,k}(t)}}\right) dx \\ &= \int_t^\infty \frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)} \ln\left(\frac{f_{n,k}(x)\bar{F}_{m,k}(t)}{\bar{F}_{n,k}(t)f_{m,k}(x)}\right) dx, \end{aligned} \tag{6.2.1}$$

where $f_{n,k}(x)$ and $f_{m,k}(x)$ are pdfs of n^{th} and m^{th} upper k-record value given by (1.10.1). Also $\bar{F}_{n,k}(t)$ and $\bar{F}_{m,k}(t)$ are corresponding survival functions.
In terms of residual measure of entropy denoted by $H(f_{n,k};t)$, it can be written as

$$H(f_{n,k}/f_{m,k};t) = -H(f_{n,k};t) - \int_t^\infty \frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)} \ln\frac{f_{m,k}(x)}{\bar{F}_{m,k}(t)} dx, \tag{6.2.2}$$

where

$$\begin{aligned} H(f_{n,k};t) &= -\int_t^\infty \frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)} \ln\frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)} dx \\ &= \ln\bar{F}_{n,k}(t) - \int_t^\infty \frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)} \ln f_{n,k}(x) dx\,. \end{aligned}$$

In another way $H(f_{n,k}/f_{m,k};t)$ can be represented as

$$H(f_{n,k}/f_{m,k};t) = \ln\frac{\bar{F}_{m,k}(t)}{\bar{F}_{n,k}(t)} + \frac{1}{\bar{F}_{n,k}(t)} \int_t^\infty f_{n,k}(x) \ln\frac{f_{n,k}(x)}{f_{m,k}(x)} dx, \ \ t>0. \tag{6.2.3}$$

When $t=0$, then (6.2.1) become

$$H(f_{n,k}/f_{m,k}) = \int_0^\infty f_{n,k}(x) \ln\frac{f_{n,k}(x)}{f_{m,k}(x)} dx,$$

the extension of K-L information measure to k-record values refer to Mosayeb and Borzadaran [81].

Next, we prove that the measure of discrimination between two k-record values of residual lifetime characterizes the distribution function of parent random variable uniquely. We derive the required result using the theorem and lemma due to Gupta and Kirmani [51] which we have stated earlier (refer to Theorem 4.5.1 and Lemma 4.5.1). The result is stated as:

Theorem 6.2.1. *Let $H(f_{n,k}/f_{m,k};t)$ is the discrimination measure between two k-record values of residual lifetime corresponding to the random variable X having distribution function $F(x)$. Then this measure characterizes the distribution function of parent random variable uniquely.*

Proof. We can write from (6.2.3)

$$H(f_{n,k}/f_{m,k};t) = \ln \bar{F}_{m,k}(t) - \ln \bar{F}_{n,k}(t) + \int_t^\infty \frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)} \ln f_{n,k}(x)dx - \int_t^\infty \frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)} \ln f_{m,k}(x)dx.$$

Taking derivative both sides with respect to t, we get

$$\begin{aligned} H'(f_{n,k}/f_{m,k};t) &= -\lambda_{F_{m,k}}(t) + \lambda_{F_{n,k}}(t) \int_t^\infty \frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)} \ln f_{n,k}(x)dx - \lambda_{F_{n,k}}(t) \ln f_{n,k}(t) \\ &\quad + \lambda_{F_{n,k}}(t) - \lambda_{F_{n,k}}(t) \int_t^\infty \frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)} \ln f_{m,k}(x)dx + \lambda_{F_{n,k}}(t) \ln f_{m,k}(t) \\ &= \lambda_{F_{n,k}}(t) - \lambda_{F_{m,k}}(t) + \lambda_{F_{n,k}}(t) \ln \frac{f_{m,k}(t)}{f_{n,k}(t)} + \lambda_{F_{n,k}}(t) \int_t^\infty \frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)} \ln \frac{f_{n,k}(x)}{f_{m,k}(x)} dx. \end{aligned}$$

After some rearrangements, we get

$$H'(f_{n,k}/f_{m,k};t) = \lambda_{F_{n,k}}(t) - \lambda_{F_{m,k}}(t) + \lambda_{F_{n,k}}(t) \ln \frac{\lambda_{F_{m,k}}(t)}{\lambda_{F_{n,k}}(t)} + \lambda_{F_{n,k}}(t) H(f_{n,k}/f_{m,k};t).$$

Taking derivative with respect to t again and using the relation

$$\lambda_{F_{n,k}}(t) = c(t)\lambda_{F_{m,k}}(t),$$

where

$$c(t) = (-\ln \bar{F}(t))^{(n-m)} \frac{\Gamma(n)\Gamma(m; -k \ln \bar{F}(t))}{\Gamma(m)\Gamma(n; -k \ln \bar{F}(t))}.$$

We get

$$H''(f_{n,k}/f_{m,k};t) = c(t) - 1 + c(t)\left(-\ln c(t) + H(f_{n,k}/f_{m,k};t)\right)\lambda'_{F_{m,k}}(t) + \lambda_{F_{m,k}}(t).$$
$$\left(c'(t) + c'(t)\left(-\ln c(t) + H(f_{n,k}/f_{m,k};t)\right) + c(t)\left(-\frac{c'(t)}{c(t)} + H'(f_{n,k}/f_{m,k};t)\right)\right). \tag{6.2.4}$$

Suppose there are two distributions F_1 and F_2 such that

$$H(f^{(1)}_{n,k}/f_{m,k};t) = H(f^{(2)}_{n,k}/f_{m,k};t) = \kappa(t), say$$

then $\forall\, t$, from (6.2.4), we get

$$\lambda'_{F^{(1)}_{m,k}}(t) = f(t, \lambda_{F^{(1)}_{m,k}}(t)) \;\; and \;\; \lambda'_{F^{(2)}_{m,k}}(t) = f(t, \lambda_{F^{(2)}_{m,k}}(t)),$$

where

$$f(t,y) = \left(\frac{\kappa''(t) - y(c'(t) - c'(t)\ln c(t) + c'(t)\kappa(t))}{c(t) - 1 - c(t)\ln c(t) + c(t)\kappa(t)}\right).$$

Using Theorem 4.5.1 and Lemma 4.5.1 , we get $\lambda'_{F^{(1)}_{m,k}}(t) = \lambda'_{F^{(2)}_{m,k}}(t),\ \forall\, t$. Also hazard rate function characterizes the distribution function uniquely. So we get the desired result.

In the next section we study some properties of the measure of discrimination between two k-record values.

6.3 Properties

- *For each $t > 0,\ H(f_{n,k}/f_{m,k};t) \geq 0$ and the equality holds iff $f_{n,k}(x) = f_{m,k}(x)$ almost everywhere.*

  Here result is obvious as for each fixed $t > 0,\ H(f_{n,k}/f_{m,k};t)$ satisfies all the properties of $H(f_{n,k}/f_{m,k})$.

- *Let X and Y be two non-negative and continuous random variables having pdfs $f(x)$ and $g(y)$ respectively. Also $Y = \phi(X)$, where ϕ is a strictly monotonically increasing and differentiable function. Then*

$$H(f_{n,k}/f_{m,k};\phi^{-1}(t)) = H(g_{n,k}/g_{m,k};t), \tag{6.3.1}$$

where letters have the usual meaning.

Proof. We can write from (6.2.3)

$$H(g_{n,k}/g_{m,k};t) = \ln\frac{\bar{G}_{m,k}(t)}{\bar{G}_{n,k}(t)} + \int_t^\infty \frac{g_{n,k}(y)}{\bar{G}_{n,k}(t)} \ln\frac{g_{n,k}(y)}{g_{m,k}(y)}\, dy.$$

Using (1.10.1), we get

$$H(g_{n,k}/g_{m,k};t) = \ln\frac{\bar{G}_{m,k}(t)}{\bar{G}_{n,k}(t)} + \frac{1}{\bar{G}_{n,k}(t)}\times$$
$$\int_t^\infty \frac{k^n}{\Gamma(n)}(-\ln\bar{G}(y))^{n-1}(\bar{G}(y))^{k-1}g(y)\ln\left(k^{n-m}\frac{\Gamma(m)}{\Gamma(n)}(-\ln\bar{G}(y))^{n-m}\right)dy.$$

Now, the transformation $Y=\phi(X)\Rightarrow G(y)=F(x)\, and\, g(y)=f(x)\phi'(x)$. Therefore the measure $H(g_{n,k}/g_{m,k};t)$ becomes

$$H(g_{n,k}/g_{m,k};t) = \ln\frac{\bar{F}_{m,k}(\phi^{-1}t)}{\bar{F}_{n,k}(\phi^{-1}t)} + \frac{1}{\bar{F}_{n,k}(\phi^{-1}(t))}\times$$
$$\int_{\phi^{-1}(t)}^\infty \frac{k^n}{\Gamma(n)}(-\ln\bar{F}(x))^{n-1}(\bar{F}(x))^{k-1}f(x)\ln\left(k^{n-m}\frac{\Gamma(m)}{\Gamma(n)}(-\ln\bar{F}(x))^{n-m}\right)dx$$
$$= H(f_{n,k}/f_{m,k};\phi^{-1}(t)).$$

- *Consider three non-negative random upper k-record values $X_{n_1,k}$, $X_{n_2,k}$ and $X_{n_3,k}$ with pdfs $f_{n_1,k}(x)$, $f_{n_2,k}(x)$ and $f_{n_3,k}(x)$ and reverse hazard rate functions $\lambda_{f_{n_1,k}}(x)$, $\lambda_{f_{n_2,k}}(x)$ and $\lambda_{f_{n_3,k}}(x)$ respectively. If*

a. $\frac{f_{n_1,k}(x)}{f_{n_3,k}(x)}$ *is increasing in x,*

b. $\lambda_{f_{n_2,k}}(x)\le\lambda_{f_{n_1,k}}(x)$, *then* $H(f_{n_1,k}/f_{n_3,k};t)\le H(f_{n_2,k}/f_{n_3,k};t)$.

Proof. Using (6.2.1), we can write

$$H(f_{n_1,k}/f_{n_3,k};t) - H(f_{n_2,k}/f_{n_3,k};t)$$
$$= \int_t^\infty \frac{f_{n_1,k}(x)}{\bar{F}_{n_1,k}(t)}\ln\left(\frac{f_{n_1,k}(x)\bar{F}_{n_3,k}(t)}{\bar{F}_{n_1,k}(t)f_{n_3,k}(x)}\right)dx - \int_t^\infty \frac{f_{n_2,k}(x)}{\bar{F}_{n_2,k}(t)}\ln\left(\frac{f_{n_2,k}(x)\bar{F}_{n_3,k}(t)}{\bar{F}_{n_2,k}(t)f_{n_3,k}(x)}\right)dx$$

$$= \int_t^\infty \frac{f_{n_1,k}(x)}{\bar{F}_{n_1,k}(t)}\ln\left(\frac{f_{n_1,k}(x)\bar{F}_{n_3,k}(t)}{\bar{F}_{n_1,k}(t)f_{n_3,k}(x)}\right)dx - \int_t^\infty \frac{f_{n_2,k}(x)}{\bar{F}_{n_2,k}(t)}\ln\left(\frac{f_{n_1,k}(x)\bar{F}_{n_3,k}(t)}{\bar{F}_{n_1,k}(t)f_{n_3,k}(x)}\right)dx -$$
$$\int_t^\infty \frac{f_{n_2,k}(x)}{\bar{F}_{n_2,k}(t)}\ln\left(\frac{f_{n_2,k}(x)\bar{F}_{n_1,k}(t)}{\bar{F}_{n_2,k}(t)f_{n_1,k}(x)}\right)dx$$

$$\leq \int_t^\infty \frac{f_{n_1,k}(x)}{\bar{F}_{n_1,k}(t)} \ln\left(\frac{f_{n_1,k}(x)\bar{F}_{n_3,k}(t)}{\bar{F}_{n_1,k}(t) f_{n_3,k}(x)}\right) dx - \int_t^\infty \frac{f_{n_2,k}(x)}{\bar{F}_{n_2,k}(t)} \ln\left(\frac{f_{n_1,k}(x)\bar{F}_{n_3,k}(t)}{\bar{F}_{n_1,k}(t) f_{n_3,k}(x)}\right) dx$$
$$= \int_t^\infty \frac{f_{n_1,k}(x)}{\bar{F}_{n_1,k}(t)} \ln\left(\frac{f_{n_1,k}(x)}{f_{n_3,k}(x)}\right) dx - \int_t^\infty \frac{f_{n_2,k}(x)}{\bar{F}_{n_2,k}(t)} \ln\left(\frac{f_{n_1,k}(x)}{f_{n_3,k}(x)}\right) dx .$$

Here we use the fact that $H(f_{n_2,k}/f_{n_1,k};t) \geq 0$. Also (b) implies that $X^{(t)}_{n_2,k}$ is stochastically larger than $X^{(t)}_{n_1,k}$, where $X^{(t)}_{n_i,k}$, $i = 1,2,3$ is a random record variable with pdf $f^{(t)}_{n_i,k} = \frac{f_{n_i,k}(x)}{\bar{F}_{n_i,k}(t)}$. Using (a) and (b), the expression is non-positive. This completes the proof. □

6.4 Measure Of Discrimination For Some Specific Distributions

First we prove the following lemma which provides a general expression for the measure (6.2.1) which can be used to find measure of discrimination for specific distributions.

Lemma 6.4.1. *Let $f_{n,k}$ and $f_{m,k}$ are the pdfs of n^{th} and m^{th} k-record values given by (1.10.1), then the measure (6.2.1) is given as*

$$H(f_{n,k}/f_{m,k};t) = \ln \frac{\bar{F}_{m,k}(t)\Gamma(m)}{\bar{F}_{n,k}(t)\Gamma(n)} + \frac{(n-m)}{\bar{F}_{n,k}(t)\Gamma(n)} \int_{-k\ln\bar{F}(t)}^\infty u^{n-1}e^{-u} \ln u \, du . \qquad (6.4.1)$$

Proof. We know from (6.2.3)

$$H(f_{n,k}/f_{m,k};t) = \ln \frac{\bar{F}_{m,k}(t)}{\bar{F}_{n,k}(t)} + \int_t^\infty \frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)} \ln\left(\frac{f_{n,k}(x)}{f_{m,k}(x)}\right) dx.$$

Using (1.10.1) in above , we get

$$H(f_{n,k}/f_{m,k};t) = \ln \frac{\bar{F}_{m,k}(t)}{\bar{F}_{n,k}(t)} + \ln\left(\frac{k^{n-m}\Gamma(m)}{\Gamma(n)}\right) + \frac{1}{\bar{F}_{n,k}(t)} \times$$
$$\int_t^\infty \frac{k^n}{\Gamma(n)} \left(-\ln\bar{F}(x)\right)^{n-1} \left(\bar{F}(x)\right)^{k-1} f(x) \ln\left(-\ln\bar{F}(x)\right)^{(n-m)} dx.$$

After substituting $-k\ln\bar{F}(x) = u$, we obtain

$$H(f_{n,k}/f_{m,k};t) = \ln \frac{\bar{F}_{m,k}(t)k^{n-m}\Gamma(m)}{\bar{F}_{n,k}(t)\Gamma(n)} + \frac{(n-m)}{\bar{F}_{n,k}(t)} \int_{-k\ln\bar{F}(t)}^\infty \frac{u^{n-1}e^{-u}}{\Gamma(n)} (\ln u - \ln k) \; du .$$

$$\Rightarrow H(f_{n,k}/f_{m,k};t) = \ln \frac{\bar{F}_{m,k}(t)\Gamma(m)}{\bar{F}_{n,k}(t)\Gamma(n)} + \frac{(n-m)}{\bar{F}_{n,k}(t)} \int_{-k\ln\bar{F}(t)}^\infty \frac{u^{n-1}e^{-u}}{\Gamma(n)} \ln u \, du .$$

This completes the proof. □

Remark 6.4.1. *In terms of expectation it can be written as*

$$H(f_{n,k}/f_{m,k};t) = \ln\frac{\bar{F}_{m,k}(t)\Gamma(m)}{\bar{F}_{n,k}(t)\Gamma(n)} + \frac{(n-m)}{\bar{F}_{n,k}(t)}E(\ln\nu^*), \tag{6.4.2}$$

where ν^* follows the gamma distribution. That is $\nu^* \sim \Gamma(n; -k\ln\bar{F}(t))$.

The above expression gives the measure of discrimination when we keep k fixed. Next, we give an expression for measure when n is fixed but k varies.

Lemma 6.4.2. *Let $f_{n,k}$ and $f_{n,l}$ are n^{th} k-record value and l-record value with survival functions $\bar{F}_{n,k}(t)$ and $\bar{F}_{n,l}(t)$ respectively. Then*

$$H(f_{n,k}/f_{n,l};t) = \ln\frac{\bar{F}_{n,l}(t)k^n}{\bar{F}_{n,k}(t)l^n} + \frac{(l-k)n}{k\bar{F}_{n,k}(t)}\bar{F}_{n+1,k}. \tag{6.4.3}$$

Proof. We know from (6.2.3)

$$H(f_{n,k}/f_{n,l};t) = \ln\frac{\bar{F}_{n,l}(t)}{\bar{F}_{n,k}(t)} + \int_t^\infty \frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)}\ln\frac{f_{n,k}(x)}{f_{n,l}(x)}dx.$$

After using (1.10.1), we get

$$H(f_{n,k}/f_{n,l};t) = \ln\frac{\bar{F}_{n,l}(t)}{\bar{F}_{n,k}(t)} + n\ln\frac{k}{l} + \frac{(k-l)}{\bar{F}_{n,k}(t)}\int_t^\infty f_{n,k}(x)\ln\bar{F}(x)dx.$$

Substituting $-k\ln\bar{F}(x) = u$ and after some rearrangements, we get

$$\begin{aligned} H(f_{n,k}/f_{n,l};t) &= \ln\left(\frac{\bar{F}_{n,l}(t)k^n}{\bar{F}_{n,k}(t)l^n}\right) - \frac{(k-l)}{k\bar{F}_{n,k}(t)}\int_{-k\ln\bar{F}(t)}^\infty \frac{u^n e^{-u}}{\Gamma(n)}du \\ &= \ln\left(\frac{\bar{F}_{n,l}(t)k^n}{\bar{F}_{n,k}(t)l^n}\right) + \frac{(l-k)n}{k\bar{F}_{n,k}(t)}\bar{F}_{n+1,k}(t). \end{aligned}$$

Hence the result. □

Example 6.4.1. *Let X be a random variable having finite range distribution with pdf $f(x) = \frac{a}{b}\left(1-\frac{x}{b}\right)^{(a-1)}$, $a>1$, $0 \le x \le b$ and survival function $F(x) = 1-\left(1-\frac{x}{b}\right)^a$.*
$\Rightarrow \quad \bar{F}_{n,k}(t) = \frac{\Gamma(n;-k\ln(1-\frac{t}{b})^a)}{\Gamma n}$ *and* $\bar{F}_{n,l}(t) = \frac{\Gamma(n;-l\ln(1-\frac{t}{b})^a)}{\Gamma n}$.
Hence from (6.4.1), we get

$$H(f_{n,k}/f_{m,k};t) = \ln\frac{\Gamma(m;-k\ln(1-\frac{t}{b})^a)}{\Gamma(n;-k\ln(1-\frac{t}{b})^a)} + \frac{(n-m)}{\Gamma(n;-k\ln(1-\frac{t}{b})^a)}E(\ln\nu^*)\,, \tag{6.4.4}$$

where $\nu^* \sim \Gamma(n; -k\ln\left(1-\frac{t}{b}\right)^a)$.

Also from (6.4.3), we get

$$H(f_{n,k}/f_{n,l};t) = \ln\frac{\Gamma(n;-l\ln(1-\frac{t}{b})^a)k^n}{\Gamma(n;-k\ln(1-\frac{t}{b})^a)l^n} + \frac{(l-k)\Gamma(n+1;-k\ln(1-\frac{t}{b})^a)}{k\Gamma(n;-k\ln(1-\frac{t}{b})^a)}. \tag{6.4.5}$$

Example 6.4.2. *For $a = 1$, the finite range distribution becomes uniform distribution with pdf $f(x) = \frac{1}{b}$ for $0 \le x \le b$. Hence putting $a = 1$ in (6.4.4) and (6.4.5), we get discrimination measures for uniform distribution as*

$$H(f_{n,k}/f_{m,k};t) = \ln\frac{\Gamma(m;-k\ln(1-\frac{t}{b}))}{\Gamma(n;-k\ln(1-\frac{t}{b}))} + \frac{(n-m)}{\Gamma(n;-k\ln(1-\frac{t}{b}))}E(\ln\nu^*),$$

where $\nu^* \sim \Gamma(n; -k\ln\left(1-\frac{t}{b}\right))$

and

$$H(f_{n,k}/f_{n,l};t) = \ln\frac{\Gamma(n;-l\ln(1-\frac{t}{b}))k^n}{\Gamma(n;-k\ln(1-\frac{t}{b}))l^n} + \frac{(l-k)\Gamma(n+1;-k\ln(1-\frac{t}{b}))}{k\Gamma(n;-k\ln(1-\frac{t}{b}))}.$$

Remark 6.4.2. *A relation among discrimination measure between two distribution with discrimination measure between their corresponding residual lifetime distribution and discrimination measure between their past lifetime distributions is given by*

$$H(f_{n,k}/f_{m,k}) = F_{n,k}(t)\bar{H}(f_{n,k}/f_{m,k};t) + F_{n,k}(t)\ln\frac{F_{n,k}(t)}{F_{m,k}(t)} + \bar{F}_{n,k}(t)H(f_{n,k}/f_{m,k};t) + \bar{F}_{n,k}(t)\ln\frac{\bar{F}_{n,k}(t)}{\bar{F}_{m,k}(t)}, \tag{6.4.6}$$

where $\bar{H}(f_{n,k}/f_{m,k};t)$ denotes the discrimination measure between two past lifetime record distributions. Since

$$H(f_{n,k}/f_{m,k}) = \int_0^\infty f_{n,k}(x)\ln\frac{f_{n,k}(x)}{f_{m,k}(x)}dx.$$

refer to Mosayeb and Borzadaran [81].

After some rearrangements, we get

$$\begin{aligned}
H(f_{n,k}/f_{m,k}) =& F_{n,k}(t)\int_0^t \frac{f_{n,k}(x)}{F_{n,k}(t)}\ln\left(\frac{f_{n,k}(x)F_{m,k}(t)}{F_{n,k}(t)f_{m,k}(x)}\right)dx + F_{n,k}(t)\ln\frac{F_{n,k}(t)}{F_{m,k}(t)} + \\
&\bar{F}_{n,k}(t)\int_t^\infty \frac{f_{n,k}(x)}{\bar{F}_{n,k}(t)}\ln\left(\frac{f_{n,k}(x)\bar{F}_{m,k}(t)}{\bar{F}_{n,k}(t)f_{m,k}(x)}\right)dx + \bar{F}_{n,k}(t)\ln\frac{\bar{F}_{n,k}(t)}{\bar{F}_{m,k}(t)} \\
=& F_{n,k}(t)\bar{H}(f_{n,k}/f_{m,k};t) + F_{n,k}(t)\ln\frac{F_{n,k}(t)}{F_{m,k}(t)} + \bar{F}_{n,k}(t)H(f_{n,k}/f_{m,k};t) + \bar{F}_{n,k}(t)\ln\frac{\bar{F}_{n,k}(t)}{\bar{F}_{m,k}(t)}.
\end{aligned}$$

6.5 Bounds To The Measure

1. *If $f(x)$ is a increasing function then $\forall\, n > m$,*

$$H(f_{n,k}/f_{m,k};t) \geq \ln \frac{\lambda_{F_{n,k}(t)}}{\lambda_{F_{m,k}(t)}}. \tag{6.5.1}$$

and inequality is reversed if $f(x)$ is a decreasing function, where $\lambda_{F_{n,k}(t)}$ and $\lambda_{F_{m,k}(t)}$ denote the hazard rates of $X_{n,k}$ and $X_{m,k}$ respectively.

Proof. We know

$$\frac{f_{n,k}(x)}{f_{m,k}(x)} = \frac{k^{(n-m)}\Gamma(m)}{\Gamma(n)} \left(-\ln \bar{F}(x)\right)^{(n-m)}. \tag{6.5.2}$$

As $f(x)$ is an increasing function, this gives $\forall\, n > m$, $\frac{f_{n,k}(x)}{f_{m,k}(x)}$ is also increasing function of x

Therefore

$$x \geq t \;\Rightarrow \frac{f_{n,k}(x)}{f_{m,k}(x)} \geq \frac{f_{n,k}(t)}{f_{m,k}(t)}.$$

From (6.2.1), we have

$$\begin{aligned} H(f_{n,k}/f_{m,k};t) &= \frac{1}{\bar{F}_{n,k}(t)} \int_t^{\infty} f_{n,k}(x) \ln \left(\frac{f_{n,k}(x)\bar{F}_{m,k}(t)}{\bar{F}_{n,k}(t) f_{m,k}(x)} \right) dx \\ &\geq \frac{1}{\bar{F}_{n,k}(t)} \int_t^{\infty} f_{n,k}(x) \ln \left(\frac{f_{n,k}(t)\bar{F}_{m,k}(t)}{\bar{F}_{n,k}(t) f_{m,k}(t)} \right) dx \\ &= \ln \frac{\lambda_{F_{n,k}(t)}}{\lambda_{F_{m,k}(t)}}. \end{aligned}$$

Also we see that (6.5.1) can be expressed in terms of survival function as follows

$$H(f_{n,k}/f_{m,k};t) \geq \ln \left(\frac{k^{(n-m)}\Gamma(m)}{\Gamma(n)} \left(-\ln \bar{F}(t)\right)^{(n-m)} \frac{\bar{F}_{m,k}(t)}{\bar{F}_{n,k}(t)} \right). \tag{6.5.3}$$

Similarly if $f(x)$ is a decreasing function. Then for $x \geq t \Rightarrow \frac{f_{n,k}(x)}{f_{m,k}(x)} \leq \frac{f_{n,k}(t)}{f_{m,k}(t)}$. Therefore we get,

$$H(f_{n,k}/f_{m,k};t) \leq \ln \frac{\lambda_{F_{n,k}(t)}}{\lambda_{F_{m,k}(t)}}.$$

□

2. *Also when $f(x)$ is increasing, then*

$$\bar{H}(f_{n,k}/f_{m,k};t) \leq \ln \frac{r_{F_{n,k}(t)}}{r_{F_{m,k}(t)}}, \tag{6.5.4}$$

where $\bar{H}(f_{n,k}/f_{m,k};t)$ represents the discrimination measure between two past lifetime distributions of records and $r_{F_{n,k}(t)}$ represents reverse hazard rate function of $X_{n,k}$ and inequality is reversed if $f(x)$ is a decreasing function.

Proof. Proof is on similar lines as above and hence omitted. □

3. *If $f(x)$ is an increasing function. Then*

$$H(f_{n,k}/f_{m,k}) \leq \ln \frac{f_{n,k}(t)}{f_{m,k}(t)} \tag{6.5.5}$$

and inequality is reversed if $f(x)$ is decreasing function.

Proof. We know from (6.4.6)

$$H(f_{n,k}/f_{m,k}) = F_{n,k}(t)\bar{H}(f_{n,k}/f_{m,k};t) + F_{n,k}(t)\ln\frac{F_{n,k}(t)}{F_{m,k}(t)} + \bar{F}_{n,k}(t)H(f_{n,k}/f_{m,k};t) + \bar{F}_{n,k}(t)\ln\frac{\bar{F}_{n,k}(t)}{\bar{F}_{m,k}(t)}.$$

Putting (6.5.2) and (6.5.4) in this, we get

$$\begin{aligned}
H(f_{n,k}/f_{m,k}) &\leq F_{n,k}(t)\ln\frac{r_{F_{n,k}}(t)}{r_{F_{m,k}}(t)} - \bar{F}_{n,k}(t)\ln\frac{\lambda_{F_{m,k}}(t)}{\lambda_{F_{n,k}}(t)} + F_{n,k}(t)\ln\frac{F_{n,k}(t)}{F_{m,k}(t)} + \bar{F}_{n,k}(t)\ln\frac{\bar{F}_{n,k}(t)}{\bar{F}_{m,k}(t)} \\
&= F_{n,k}(t)\ln\left(\frac{r_{F_{n,k}}(t)F_{n,k}(t)}{r_{F_{m,k}}(t)F_{m,k}(t)}\right) + \bar{F}_{n,k}(t)\ln\left(\frac{\bar{F}_{n,k}(t)\lambda_{F_{n,k}}(t)}{\bar{F}_{m,k}(t)\lambda_{F_{m,k}}(t)}\right) \\
&= F_{n,k}(t)\ln\frac{f_{n,k}(t)}{f_{m,k}(t)} + \bar{F}_{n,k}(t)\ln\frac{f_{n,k}(t)}{f_{m,k}(t)} \\
&= \left(F_{n,k}(t) + \bar{F}_{n,k}(t)\right)\ln\frac{f_{n,k}(t)}{f_{m,k}(t)}.
\end{aligned}$$

$$\Rightarrow H(f_{n,k}/f_{m,k}) \leq \ln\frac{f_{n,k}(t)}{f_{m,k}(t)}.$$

Also we see that (6.5.5) can be expressed in terms of survival function as follows

$$H(f_{n,k}/f_{m,k}) \leq \ln\left(\frac{k^{(n-m)}\Gamma(m)}{\Gamma(n)}\left(-\ln\bar{F}(t)\right)^{(n-m)}\right). \tag{6.5.6}$$

□

Now, we give some examples based on above bounds.

Example 6.5.1. *Let X be a exponentially distributed random variable with pdf* $f(x) = ae^{-ax}$ *and the survival function* $\bar{F}(x) = e^{-ax}$, $a \geq 0$. *Here* $f(x)$ *is a decreasing function. Therefore from (6.5.3) and using (6.5.6) for decreasing function, we get*

$$H(f_{n,k}/f_{m,k};t) \leq \ln\left(\frac{k^{(n-m)}\Gamma(m)}{\Gamma(n)}(-\ln\bar{F}(t))^{(n-m)}\frac{\bar{F}_{m,k}(t)}{\bar{F}_{n,k}(t)}\right)$$
$$= \ln\left(\frac{(kat)^{(n-m)}\Gamma(m;-k\ln\bar{F}(t))}{\Gamma(n;-k\ln\bar{F}(t))}\right)$$

and

$$H(f_{n,k}/f_{m,k}) \geq \ln\left(\frac{(kat)^{(n-m)}\Gamma(m)}{\Gamma(n)})\right).$$

Example 6.5.2. *Let X be an absolutely continuous random variable with support* $(0,1]$ *with pdf*

$$f(x) = 2x, \;\; if x\,\varepsilon\;(0,1]$$

and survival function $\bar{F}(x) = 1-x^2$. *It is an increasing function on* $(0,1]$ *and hence using (6.5.3), we get*

$$H(f_{n,k}/f_{m,k};t) \geq \ln\left(\frac{k^{(n-m)}\Gamma(m)}{\Gamma(n)}(-\ln\bar{F}(t))^{(n-m)}\frac{\bar{F}_{m,k}(t)}{\bar{F}_{n,k}(t)}\right)$$
$$= \ln\left(\left(-k\ln(1-t^2)\right)^{(n-m)}\frac{\Gamma(m;-k\ln(1-t^2))}{\Gamma(n;-k\ln(1-t^2))}\right).$$

and also from (6.5.6)

$$H(f_{n,k}/f_{m,k}) \leq \ln\left(\left(-k\ln(1-t^2)\right)^{(n-m)}\frac{\Gamma(m)}{\Gamma(n)}\right).$$

6.6 Conclusion

Considering the importance of discrimination measure, in this chapter we have proposed a measure of discrimination between two k-record value distributions and a measure of discrimination between k-record value and l-record value distributions of residual lifetime corresponding to the random variable X and studied a characterization result for that. Also we have studied some of the properties of the proposed measure like non negativity and effect of monotone transformation on it. Also some upper and lower bounds to the discrimination measure have been obtained.

Chapter 7

Generalized Entropy Measure For k-Record Values

7.1 Introduction

Shannon [104] entropy measure plays a crucial role in information theory. Though this measure is a focal part of information theory, yet the idea of information is so rich that no single definition have the capacity to measure the information legitimately. Hence many researchers presented the parametric group of entropies as a mathematical generalization of Shannon's entropy. These entropies are functions of some parameters and tend to Shannon entropy when these parameters approach their limiting values. A huge literature is devoted to the characterizations, generalizations and applications of the Shannon information measure, refer to Cover and Thomas [27], Aczel and Daroczy [3], Wells [123] etc. A two parametric generalization of Shannon entropy measure has been given by Verma [122]. He introduced the generalization of order α and type β of the entropy as

$$H_{\alpha}^{\beta}(X) = \frac{1}{\beta-\alpha}\log\left[\int_0^{\infty} f^{\alpha+\beta-1}(x)dx\right];\ \beta-1<\alpha<\beta,\ \beta\geq 1, \tag{7.1.1}$$

The result of this chapter has been communicated in a research paper under the title **On Generalized Information Measure of Order (α,β) and k-Record Statistics.**

where

$$\lim_{\beta=1} H_{\alpha}^{\beta}(X) = H_{\alpha}(X) = \frac{1}{1-\alpha}\log\left[\int_0^{\infty} f^{\alpha}(x)dx\right], \tag{7.1.2}$$

is the Renyi entropy [99] and

$$\lim_{\beta=1,\alpha\to 1} H_{\alpha}^{\beta}(X) = -\int_0^{\infty} f(x)\log f(x)dx,$$

is the Shannon entropy [104].

The Shannon's entropy and their generalizations on order statistics and record data have been studied by several authors, refer to Ebrahimi et al. [42], Baratpour et al. [19, 20], Madadi and Tata [76], Asha and Chacko [14] and Kumar [71]. Some authors have studied the various characterizations of distribution function F based on the properties of order statistics and record values, refer to Zahedi and Shakil [126], and Raqab and Awad [98]. The theory of k^{th} records is still developing. An interesting question is whether we can determine the generalized version of measure of information contained in a sequence of k-record values from a sequence of iid random variables. Taking idea from this, here in this chapter, we study the generalized entropy measure (7.1.1) for k-record values and further we study its dynamic version.

The chapter is organized as follows: The generalized entropy of order α and type β of k-record values associated with some lifetime distributions is presented in Section 7.2. In Section 7.3, we study characterization result for the generalized entropy measure. Generalized residual entropy of order α and type β for k-record values has been studied in Section 7.4 and its characterization result in Section 7.5. An upper bound to this residual entropy of order (α, β) has been derived in Section 7.6 and chapter is concluded in Section 7.7.

7.2 Generalized Entropy For k-Record values Obtained For Some Specific Distributions

Corresponding to the generalized entropy measure (7.1.1), we propose the generalized entropy measure for the k-record values as

$$H_{\alpha}^{\beta}(X_{n,k}) = \frac{1}{\beta-\alpha}\log\left\{\int_0^{\infty}(f_{n,k}(x))^{\alpha+\beta-1}dx\right\},\ \beta-1<\alpha<\beta,\ \beta\geq 1. \tag{7.2.1}$$

Now, we use the probability integral transformation $U = F(X)$, where the distribution of U is the standard uniform distribution, for the generalized entropy measure (7.1.1). The probability integral transformation provides the following useful representation for the random variable X

$$H_{\alpha}^{\beta}(X) = \frac{1}{\beta-\alpha}\log\left\{\int_0^{1} f^{(\alpha+\beta-2)}(F^{-1}(u)du\right\}. \tag{7.2.2}$$

Next we prove the following result.

Lemma 7.2.1. *The entropy measure (7.2.1) of the n^{th} upper* k*-record values $X_{n,k}$ can be expressed as*

$$\begin{aligned} H_{\alpha}^{\beta}(X_{n,k}) &= \frac{1}{\beta-\alpha}\log\left\{\frac{k^{n(\alpha+\beta-1)}\,\Gamma((n-1)(\alpha+\beta-1)+1)}{((k-1)(\alpha+\beta-1)+1)^{(n-1)(\alpha+\beta-1)+1}\,(\Gamma(n))^{(\alpha+\beta-1)}}\times\right.\\ &\quad \left. E\{f^{\alpha+\beta-2}(F^{-1}(1-e^{-u}))\}\right\}, \end{aligned} \tag{7.2.3}$$

where u follows gamma distribution with parameter $((n-1)(\alpha+\beta-1)+1)$ and $((k-1)(\alpha+\beta-1)+1)$ and E denotes the expectation.

Proof. Using (1.10.1), the entropy measure (7.2.1) can be rewritten as

$$\begin{aligned} H_{\alpha}^{\beta}(X_{n,k}) &= \frac{1}{\beta-\alpha}\log\left\{\frac{k^{n(\alpha+\beta-1)}}{(\Gamma(n))^{(\alpha+\beta-1)}}\times\right. \\ &\quad \left.\int_0^{\infty}(-\log\bar{F}(x))^{(n-1)(\alpha+\beta-1)}\,(\bar{F}(x))^{(k-1)(\alpha+\beta-1)}f^{(\alpha+\beta-1)}(x)dx\right\}. \end{aligned} \tag{7.2.4}$$

Substituting $-\log\bar{F}(x) = u$, and hence $x = F^{-1}(1-e^{-u})$, we have

$$\begin{aligned} H_{\alpha}^{\beta}(X_{n,k}) &= \frac{1}{\beta-\alpha}\log\left\{\frac{k^{n(\alpha+\beta-1)}}{(\Gamma(n))^{(\alpha+\beta-1)}}\times\right. \\ &\quad \left.\int_0^{\infty}u^{(n-1)(\alpha+\beta-1)}e^{-u((k-1)(\alpha+\beta-1)+1)}\left\{f^{\alpha+\beta-2}(F^{-1}(1-e^{-u}))\right\}du\right\}. \end{aligned}$$

It can be rewritten as

$$H_{\alpha}^{\beta}(X_{n,k}) = \frac{1}{\beta-\alpha}\log\left\{\frac{k^{n(\alpha+\beta-1)}\,\Gamma((n-1)(\alpha+\beta-1)+1)}{((k-1)(\alpha+\beta-1)+1)^{(n-1)(\alpha+\beta-1)+1}\,(\Gamma(n))^{(\alpha+\beta-1)}}E\{f^{(\alpha+\beta)-2}(F^{-1}(1-e^{-u}))\}\right\}$$

So, the result follows. □

For some specific univariate continuous distributions, the expression (7.2.3) is evaluated, refer to Table 7.1, and the generalized entropy (7.2.2) for the parent distribution is provided in Table 7.2 which can be obtained from Table 7.1 by taking $n = k = 1$.

Distribution Function	Generalized Entropy $H_{\alpha}^{\beta}(X_{n,k})$
Uniform, $X \sim U(a,b)$	$\frac{1}{\beta-\alpha}\log\left\{\frac{k^{n(\alpha+\beta-1)}\Gamma[(n-1)(\alpha+\beta-1)+1]}{(b-a)^{\alpha+\beta-2}(\Gamma(n))^{\alpha+\beta-1}}\right\}$
Exponential, $X \sim \exp(\theta)$	$\frac{1}{\beta-\alpha}\log\left\{\frac{k^{n(\alpha+\beta-1)}\theta^{(\alpha+\beta-2)}\Gamma((n-1)(\alpha+\beta-1)+1)}{(\Gamma(n))^{(\alpha+\beta-1)}\{k(\alpha+\beta-1)\}^{(n-1)(\alpha+\beta-1)+1}}\right\}$
Pareto, $X \sim P(a)$	$\frac{1}{\beta-\alpha}\log\left\{\frac{\Gamma[(n-1)(\alpha+\beta-1)+1]k^{n(\alpha+\beta-1)}a^{\alpha+\beta-2}}{\left(\Gamma(n)\right)^{\alpha+\beta-1}\{k(\alpha+\beta-1)+\frac{\alpha+\beta-2}{a}\}^{(n-1)(\alpha+\beta-1)+1}}\right\}$
Weibull, $X \sim W(a,b)$	$\frac{1}{\beta-\alpha}\log\left\{\frac{k^{n(\alpha+\beta-1)}\left(ba^{\frac{1}{b}}\right)^{\alpha+\beta-2}\Gamma(n(\alpha+\beta-1)-\frac{\alpha+\beta-2}{b})}{(\Gamma(n))^{\alpha+\beta-1}\{k(\alpha+\beta-1)\}^{(n(\alpha+\beta-1)-\frac{\alpha+\beta-2}{b})}}\right\}$
Finite Range	$\frac{1}{\beta-\alpha}\log\left\{\frac{\Gamma((n-1)(\alpha+\beta-1)+1)}{\{ka(\alpha+\beta-1)-(\alpha+\beta-2)\}^{((n-1)(\alpha+\beta-1)+1)}}\right\} + \frac{1}{(\beta-\alpha)}\log\left\{\frac{(ka)^{n(\alpha+\beta-1)}}{b^{\alpha+\beta-2}(\Gamma(n))^{\alpha+\beta-1}}\right\}$

Table 7.1: Generalized entropy of k-record $H_{\alpha}^{\beta}(X_{n,k})$ for various lifetime distributions.

Distribution Function	Generalized Entropy $H_{\alpha}^{\beta}(X)$
Uniform, $X \sim U(a,b)$	$= \left(\frac{2-\alpha-\beta}{\beta-\alpha}\right)\log(b-a)$
Exponential, $X \sim \exp(\theta)$	$= \frac{1}{(\beta-\alpha)}\log\left(\frac{\theta^{\alpha+\beta-2}}{\alpha+\beta-1}\right)$
Pareto, $X \sim P(a)$	$= \frac{1}{(\beta-\alpha)}\log\left\{\frac{a^{\alpha+\beta-1}}{((\alpha+\beta-1)a+\alpha+\beta-2)}\right\}$
Weibull, $X \sim W(a,b)$	$= \frac{1}{(\beta-\alpha)}\log\left\{\frac{\left(ba^{\frac{1}{b}}\right)^{\alpha+\beta-2}\Gamma((\alpha+\beta-1)-\frac{\alpha+\beta-2}{b})}{(\alpha+\beta-1)^{((\alpha+\beta-1)-\frac{\alpha+\beta-2}{b})}}\right\}$
Finite Range , $X \sim FR(a,b)$	$= \frac{1}{(\beta-\alpha)}\log\left\{\frac{(a)^{(\alpha+\beta-1)}(b)^{(2-\alpha-\beta)}}{(\alpha+\beta-1)a-(\alpha+\beta-2)}\right\}$

Table 7.2: Generalized entropy $H_{\alpha}^{\beta}(X)$ for various lifetime distributions.

7.3 Characterization Result

Next, we show that distribution function F can be uniquely specified up to a location change by the generalized entropy of order (α,β) for n^{th} upper $k-$record values. For this, we will use the lemma due to Goffman and Pedrick [48] which we have stated earlier (refer to Lemma 3.2.1). The characterization result is stated as follows.

Theorem 7.3.1. *Let X and Y be two non-negative random variables having common support. Let $H_{\alpha}^{\beta}(X_{n,k}) < \infty$ and $H_{\alpha}^{\beta}(Y_{n,k}) < \infty$ be their generalized entropies of n^{th} upper k-record values respectively. Then F and G belong to the same location family of distributions, if and only if*

$$H_{\alpha}^{\beta}(X_{n,k}) = H_{\alpha}^{\beta}(Y_{n,k}) \;,\; \forall\, n,k \geq 1.$$

The proof of this follows on the same lines as in case of characteriztion result for Varma entropy for record values given by Kumar [71], and hence is omitted.

7.4 Generalized Residual Entropy For k-Record Values

The lifetime of a component has been considered as a prime variable of interest in many field such as reliability, survival analysis and economics, business etc. In such cases, the information measures are functions of time, thus as such they are dynamic. In reliability theory and survival analysis, X with density function $f(x)$ usually denotes a duration such as the lifetime of a component. The residual lifetime of the system when it is still operating at time t is $X_t = (X - t|X > t)$ and has the probability density $f(x;t) = \frac{f(x)}{\bar{F}(t)}$, $x \geq t > 0$. Ebrahimi [40] proposed the entropy of the residual lifetime X_t as

$$H(X;t) = -\int_t^\infty \frac{f(x)}{\bar{F}(t)} \log \frac{f(x)}{\bar{F}(t)} dt\ ,\ \ t > 0. \tag{7.4.1}$$

Similar results in case of a generalized residual entropy have been derived by Belzunce et al. [21] and Nanda and Paul [88]. Baig and Dar [15] have proposed the dynamic residual entropy of order α and type β as

$$H_\alpha^\beta(X;t) = \frac{1}{\beta - \alpha} \log \left\{ \frac{\int_t^\infty f^{\alpha+\beta-1}(x)dx}{\bar{F}^{\alpha+\beta-1}(t)} \right\};\ \ \beta - 1 < \alpha < \beta,\ \beta \geq 1, \tag{7.4.2}$$

and studied its properties. For $\beta = 1$, (7.4.2) reduces to residual Renyi entropy given by

$$H_\alpha(X;t) = \frac{1}{1-\alpha} \log \left\{ \frac{\int_t^\infty f^\alpha(x)dx}{\bar{F}^\alpha(t)} \right\}\ ,$$

for more details refer to Abraham and Sankaran [2]. In actuarial science, generalized entropy given in (7.4.2) can be presented as the pre-payment entropy of claims (losses) with a deductible t.

The role of residual entropy as a measure of uncertainty in order statistics and record value has been studied by many researchers refer to, Zarezadeh and Asadi [127]. Madadi and Tata [77] generalized the results for Shannon information measure to k-records. Now, analogous to the measure (7.4.2), we propose the generalized residual entropy of order α and type β for n^{th} upper k-record value as

$$H_\alpha^\beta(X_{n,k};t) = \frac{1}{\beta - \alpha} \log \left[\frac{\int_t^\infty f_{n,k}^{\alpha+\beta-1}(x)dx}{\bar{F}_{n,k}^{\alpha+\beta-1}(t)} \right];\ \ \beta - 1 < \alpha < \beta,\ \beta \geq 1.$$

Next, we derive a simplified expression for generalized residual entropy of order α and type β for the n^{th} upper k-record value. Here we use the notation $X \sim \Gamma_t(n,\lambda)$ to indicate that X has a truncated gamma distribution with density function

$$f(x) = \frac{\lambda^n}{\Gamma(n;t)} x^{n-1} e^{-\lambda x},\ \ \ x > t > 0,$$

where $\Gamma(n;t)$ is the incomplete gamma function defined as $\Gamma(n;t)=\int_t^\infty x^{n-1}e^{-x}dx$, $n\,,\,\lambda>\,0$.

Theorem 7.4.1. *Let $X_{n,k},\ n,k>1$ be a sequence of record random variables with parent distribution $F(x)$ and density function $f(x)$. Here $F^{-1}(.)$ denotes the quantile function. Then the generalized residual entropy (7.4.2) of n^{th} upper k-record value can be expressed as*

$$H_\alpha^\beta(X_{n,k};t) = \frac{1}{\beta-\alpha}\log\Big\{\frac{k^{n(\alpha+\beta-1)}\,\Gamma((n-1)(\alpha+\beta-1)+1;-k\log\bar F(t))}{((k-1)(\alpha+\beta-1)+1)^{(n-1)(\alpha+\beta-1)+1}(\Gamma(n;-k\log\bar F(t)))^{\alpha+\beta-1}}\times E\{f^{\alpha+\beta-2}(F^{-1}(1-e^{-V_z}))\}\Big\}. \qquad (7.4.3)$$

where $z=-k\log\bar F(t)$ and $V_z\sim\Gamma\{(n-1)(\alpha+\beta-1)+1;-k\log\bar F(t)\}$ and E is the expectation.

Proof. Consider

$$H_\alpha^\beta(X_{n,k};t)=\frac{1}{\beta-\alpha}\log\left[\frac{\int_t^\infty f_{n,k}^{\alpha+\beta-1}(x)dx}{\bar F_{n,k}^{\alpha+\beta-1}(t)}\right];\ \beta-1<\alpha<\beta,\ \beta\ge 1$$

$$=\frac{1}{\beta-\alpha}\log\left\{\frac{k^{n(\alpha+\beta-1)}\int_t^\infty(-\log\bar F(x))^{(n-1)(\alpha+\beta-1)}\,(\bar F(x))^{(k-1)(\alpha+\beta-1)}f^{\alpha+\beta-1}(x)dx}{(\Gamma(n;-k\log\bar F(t)))^{\alpha+\beta-1}}\right\}.$$

Substituting $-k\log\bar F(x)=u$ and $x=F^{-1}(1-e^{-\frac{u}{k}})$, we get

$$H_\alpha^\beta(X_{n,k};t) = \frac{1}{\beta-\alpha}\log\Big\{\frac{k^{n(\alpha+\beta-1)}}{k^{(n-1)(\alpha+\beta-1)+1}(\Gamma(n;-k\log\bar F(t)))^{\alpha+\beta-1}}\times \int_{-k\log\bar F(t)}^\infty u^{(n-1)(\alpha+\beta-1)}e^{-u(\frac{(k-1)(\alpha+\beta-1)+1}{k})}\Big\{f^{\alpha+\beta-2}(F^{-1}(1-e^{-\frac{u}{k}}))\Big\}du\Big\}$$

rewriting it, we get

$$H_\alpha^\beta(X_{n,k};t) = \frac{1}{\beta-\alpha}\log\Big\{\frac{k^{n(\alpha+\beta-1)}\,\Gamma((n-1)(\alpha+\beta-1)+1;-k\log\bar F(t))}{((k-1)(\alpha+\beta-1)+1)^{(n-1)(\alpha+\beta-1)+1}(\Gamma(n;-k\log\bar F(t)))^{\alpha+\beta-1}}\times E\{f^{\alpha+\beta-2}(F^{-1}(1-e^{-V_z}))\}\Big\}.$$

Hence, the result follows. □

Example 7.4.1. *Let $U_{n,k}^*$ be the n^{th} upper k-record value for a sequence of observations* *from uniform distribution on* $(0,1)$. *Then*

$$H_{\alpha}^{\beta}(U_{n,k}^{*};t)=\frac{1}{\beta-\alpha}\log\left\{\frac{k^{n(\alpha+\beta-1)}\Gamma((n-1)(\alpha+\beta-1)+1;-k\log(1-t))}{((k-1)(\alpha+\beta-1)+1)^{(n-1)(\alpha+\beta-1)+1}\left(\Gamma(n;-k\log(1-t))\right)^{\alpha+\beta-1}}\right\}. \tag{7.4.4}$$

Example 7.4.2. *Let* $\bar{U}_{n,k}$ *be the* n^{th} *upper k-record values for a sequence of observations from standard exponential distribution. Then*

$$H_{\alpha}^{\beta}(\bar{U}_{n,k};t)=\frac{1}{\beta-\alpha}\log\left\{\frac{k^{n(\alpha+\beta-1)}}{(k(\alpha+\beta-1)+1)^{(n-1)(\alpha+\beta-1)+1}}\frac{\Gamma((n-1)(\alpha+\beta-1)+1;kt)}{(\Gamma(n;kt))^{(\alpha+\beta-1)}}\right\}. \tag{7.4.5}$$

Example 7.4.3. *For Weibull random variable X with pdf is*

$$f(x)=abx^{b-1}e^{-ax^{b}},\ \ a,\ b>0,\ x>\ 0$$

where a and b are parameters. The survival function is

$$\bar{F}(x)=1-F(x)=e^{-ax^{b}}\ .$$

Substituting $-k\log\bar{F}(x)=u$, *we observe that* $x=F^{-1}(1-e^{-\frac{u}{k}})=(\frac{u}{ak})^{\frac{1}{b}}$ *and for computing* $H_{\alpha}(X_{n,k};t)$, *we have*

$$f^{\alpha+\beta-2}\left(F^{-1}(1-e^{-\frac{u}{k}})\right)=\frac{\{b(ak)^{\frac{1}{b}}\}^{\alpha+\beta-2}u^{\frac{(\alpha+\beta-2)(b-1)}{b}}e^{-\frac{u}{k}(\alpha+\beta-2)}}{k^{\alpha+\beta-2}}.$$

Therefore

$$H_{\alpha}^{\beta}(X_{n,k};t)=\frac{1}{\beta-\alpha}\log\left\{\frac{\left(b(ak)^{\frac{1}{b}}\right)^{\alpha+\beta-2}\Gamma\left(n(\alpha+\beta-1)-\frac{\alpha+\beta-2}{b};kat^{b}\right)}{\left(\Gamma(n;kat^{b})\right)^{\alpha+\beta-1}(\alpha+\beta-1)^{n(\alpha+\beta-1)-\frac{\alpha+\beta-2}{b}}}\right\}.$$

7.5 Characterization Result For Generalized Residual Entropy

In this section, we prove that like generalized entropy for k-record values, the residual entropy measure also characterizes the distribution function uniquely. We derive the characterization result for the generalized residual entropy of the n^{th} k-record value using the sufficient condition for the uniqueness of the solution of initial value differential equations. Consider the problem of finding a sufficient condition for the unique solution of the initial value problem (IVP)

$$\frac{dy}{dx} = f(x,y),\ y(x_0) = y_0 \tag{7.5.1}$$

where f is a function of two variables whose domain is a region $D \subset R^2$, (x_0, y_0) is a point in D and y is the unknown function. By the solution of (7.5.1), we find a function which satisfies the following conditions: (i) ϕ is differentiable on I, (ii) the growth of ϕ lies in D, (iii)$\phi(x_0) = y_0$ and (iv) $\phi'(x) = f(x, \phi(x))$, for all $x\varepsilon I$. For that we use the theorem and lemma due to Gupta and Kirmani [51] (refer to Theorem 4.5.1 and Lemma 4.5.1) which will help in proving our characterization result. The result is stated as:

Theorem 7.5.1. *Let X be a non-negative continuous random variable with distribution function F(x). Let generalized residual entropy measure of the corresponding n^{th} k-record value denoted by $H_\alpha^\beta(X_{n,k};t)$ be finite for all $t \geq 0$. Then $H_\alpha^\beta(X_{n,k};t)$ characterizes the distribution.*

Proof. Suppose for two distributions F_1 and F_2 such that

$$H_\alpha^\beta(X_{1n,k};t) = H_\alpha^\beta(X_{2n,k};t)\ \ \forall t \geq 0. \tag{7.5.2}$$

We know that

$$\begin{aligned} H_\alpha^\beta(X_{n,k};t) &= \frac{1}{\beta-\alpha}\log\left[\frac{\int_t^\infty f_{n,k}^{\alpha+\beta-1}(x)dx}{\bar{F}_{n,k}^{\alpha+\beta-1}(t)}\right];\ \beta-1<\alpha<\beta,\ \beta\geq 1 \\ &= \frac{1}{\beta-\alpha}\log\int_t^\infty f_{n,k}^{\alpha+\beta-1}(x)dx - \frac{1}{\beta-\alpha}\log\bar{F}_{n,k}^{\alpha+\beta-1}(t). \end{aligned}$$

Taking derivative with respect to t both sides, we get

$$(\beta-\alpha)H_\alpha^\beta(X_{n,k};t)H_\alpha^{\beta\prime}(X_{n,k};t) = [(\alpha+\beta-1)H_\alpha^\beta(X_{n,k};t) - e^{(\alpha-\beta)}]\lambda_{F_{n,k}}(t).$$

Here $\lambda_{F_{n,k}}(t)$ denotes the hazard rate function for n^{th} k-record value corresponding to the distribution function F. Differentiating above equation with respect to t again and further simplifying, we get

$$\lambda'_{F_{n,k}}(t) = \frac{(\beta-\alpha)[-H_\alpha^{\beta\prime\prime}(X_{n,k};t) + \lambda_{F_{n,k}}^{(\alpha+\beta-1)}(t)e^{(\alpha-\beta)H_\alpha^\beta(X_{n,k};t)}H_\alpha^{\beta\prime}(X_{n,k};t)]}{(\alpha+\beta-1)(-1+e^{(\alpha-\beta)H_\alpha^\beta(X_{n,k};t)})\lambda_{F_{n,k}}^{(\alpha+\beta-2)}(t)}. \tag{7.5.3}$$

Suppose that

$$H^{\beta}_{\alpha}(X_{1n,k};t) = H^{\beta}_{\alpha}(X_{2n,k};t) = c(t), \ \ say.$$

Then for all $t \geq 0$, from (7.5.3) we get

$$\lambda'_{F_{1n,k}}(t) = \psi(t, \lambda_{F_{1n,k}}(t)), \quad \lambda'_{F_{2n,k}}(t) = \psi(t, \lambda_{F_{2n,k}}(t)),$$

where

$$\psi(t,y) = \frac{(\beta-\alpha)[-c''(t) + y^{(\alpha+\beta-1)}e^{(\alpha-\beta)c(t)}c'(t)]}{(\alpha+\beta-1)(-1+e^{(\alpha-\beta)c(t)})y^{(\alpha+\beta-2)}}.$$

Using Theorem 4.5.1 and Lemma 4.5.1, we get, $\lambda_{F_{1n,k}}(t) = \lambda_{F_{2n,k}}(t)$, $\forall t$. Since the hazard rate function characterizes the distribution function uniquely, thus we get the desired result. □

7.6 Bounds To Generalized Residual Entropy Of Upper k-Record Values

Bounds to the Renyi entropy of upper record values has been considered by Zarezadeh and Asadi [127]. Here, we introduce some bounds to the generalized residual entropy of order α and type β of k-record statistic.

Theorem 7.6.1. *Let X be a non-negative continuous random variable with hazard rate function $\lambda_F(x) = \frac{f(x)}{\bar{F}(x)}$. Also let $H^{\beta}_{\alpha}(X;t)$ and $H^{\beta}_{\alpha}(X_{n,k};t)$ denote the generalized residual entropies of X and $X_{n,k}$, respectively. Also $H^{\beta}_{\alpha}(X;t)$ is finite and $m_n = \max\{(\alpha+\beta-1)(n-1), -k\log\bar{F}(t)\}$, then for $\alpha > \beta$*

$$\begin{aligned} H_{\alpha}(X_{n,k};t) \quad >(<) \quad & \frac{\alpha+\beta-1}{\alpha-\beta}\log\Gamma(n, -k\log\bar{F}(t)) - \frac{1}{\alpha-\beta}. \\ & \log\left(\frac{k^{n(\alpha+\beta-1)}}{((k-1)(\alpha+\beta-1)+1)^{(n-1)(\alpha+\beta-1)+1}}\right) \\ & -\frac{1}{\alpha-\beta}\{(\alpha+\beta-1)(n-1)\log(m_n) - m_n\} + S(t). \end{aligned} \tag{7.6.1}$$

where $S(t) = -\frac{1}{\alpha-\beta}\log\int_t^{\infty} k\lambda_F(y) f^{\alpha+\beta-2}(y)dy.$

Proof. Here m_n is the mode of Gamma distribution with density function

$$M_n = f(m_n) = \frac{m_n^{(\alpha+\beta-1)(n-1)} e^{-m_n}}{\Gamma((\alpha+\beta-1)(n-1)+1, -k\log\bar{F}(t))}.$$

From (7.2.3) we have

$$H_\alpha^\beta(X_{n,k}) = \frac{1}{\beta-\alpha}\log\Big\{\frac{k^{n(\alpha+\beta-1)}\,\Gamma((n-1)(\alpha+\beta-1)+1;-k\log\bar{F}(t))}{((k-1)(\alpha+\beta-1)+1)^{(n-1)(\alpha+\beta-1)+1}(\Gamma(n;-k\log\bar{F}(t)))^{\alpha+\beta-1}} \times E\{f^{\alpha+\beta-2}(F^{-1}(1-e^{-V_z}))\}\Big\}.$$

Now, we write for $\alpha > \beta$

$$\begin{aligned}
&-\frac{1}{\alpha-\beta}\log E\{f^{\alpha+\beta-2}\{F^{-1}(1-e^{-V_z})\}\\
&=-\frac{1}{\alpha-\beta}\log\int_{-k\log\bar{F}(t)}^{\infty}\frac{v^{(\alpha+\beta-1)(n-1)}e^{-v}}{\Gamma((\alpha+\beta-1)(n-1)+1,-k\log\bar{F}(t))}\{f^{\alpha+\beta-2}\{F^{-1}(1-e^{-\frac{v}{k}})\}dv,\\
&>-\frac{1}{\alpha-\beta}\log f(m_n)-\frac{1}{\alpha-\beta}\log\int_{-k\log\bar{F}(t)}^{\infty}\{f^{\alpha+\beta-2}\{F^{-1}(1-e^{-\frac{v}{k}})\}dv.
\end{aligned}$$

Using the transformation $y = F^{-1}(1-e^{-\frac{v}{k}})$, we obtain

$$-\frac{1}{\alpha-\beta}\log E\{f^{\alpha+\beta-2}\{F^{-1}(1-e^{-V_z})\} = -\frac{1}{\alpha-\beta}\log M_n - \frac{1}{\alpha-\beta}\log\int_t^\infty k\lambda_F(y)f^{\alpha+\beta-2}(y)dy.$$

Therefore

$$\begin{aligned}
H_\alpha(X_{n,k};t) \;>(<)\; & \frac{\alpha+\beta-1}{\alpha-\beta}\log\Gamma(n,-k\log\bar{F}(t))\\
&-\frac{1}{\alpha-\beta}\log\left(\frac{k^{n(\alpha+\beta-1)}}{((k-1)(\alpha+\beta-1)+1)^{(n-1)(\alpha+\beta-1)+1}}\right)\\
&-\frac{1}{\alpha-\beta}\{(\alpha+\beta-1)(n-1)\log(m_n)-m_n\}+S(t).
\end{aligned}$$

Hence proved.

Remark 7.6.1. *By putting $k = 1$, the results for usual records can be obtained as special case.*

7.7 Conclusion

The two parameters generalized entropies plays a vital role as a measure of complexity and uncertainty in different areas such as physics, electronics and engineering to describe many chaotic systems. Using probability integral transformation we have studied the Varma entropy and generalized residual entropy based on k-record values. The Varma entropy measure of k-record value distributions associated with various distributions viz. uniform, exponential, weibull, pareto and finite range distributions which are commonly used in the reliability modeling, has been discussed. Also we have studied some upper and lower bounds to the generalized residual entropy measure.

Summary And Further Scope

Here we summarize the work which we have presented in this thesis and also provide some scope of further investigations which can be performed on the basis of the results reported.

Summary Of The Work Reported

Shannon [104] entropy is at the focus in information theory by providing average uncertainty of a random variable X having probability density function $f(.)$. But this measure has been found inappropriate in determining the uncertainty about the remaining lifetime $[X, X \geq t]$ of a system or a component which has already worked up to time t. Thus the concept of residual measures have been developed in the literature, refer to Ebrahimi [36]. He introduced the residual entropy measure for finding the uncertainty about the remaining lifetime. Dual to the measure proposed by Ebrahimi [36], Dicrescenzo and Longobardi [31] defined the measure of uncertainty for past lifetime for a system which has been found dead at any time 't'. In our study in Chapter 2, we have considered a measure of past entropy for n^{th} upper k-record value. We have discussed some basic properties of the proposed measure.

Kerridge [63] inaccuracy measure is one of the generalization of Shannon entropy. The inaccuracy measure has various applications in different areas of science and technology such as statistical inference, estimation and coding theory. In Chapter 3, we have explored the concept of inaccuracy measure for record values. We have introduced a measure of inaccuracy between distributions of the n^{th} record value and parent random variable and discuss some properties of it. Also keeping in mind the importance of k- record values, in Chapter 4, we have also proposed and studied a measure of inaccuracy based on Kerridge measure of inaccuracy for k-record values.

To overcome some of the limitations of Shannon's entropy measure, a new measure of uncertainty was developed by Rao et al. [95], known as cumulative residual entropy in which the probability density function $f(x)$ has been replaced with the survival function $\bar{F}(x)$ of the random variable X, This measure of uncertainty is particularly appropriate to describe the information in problems connected with the ageing properties of reliability theory based on the mean residual life function. Taneja and Kumar [116] extended the concept of CRE to cumulative residual inaccuracy and then to dynamic cumulative inaccuracy and also studied some of its properties. We have provided an extension of cumulative residual inaccuracy as suggested by Taneja and Kumar [116] to k-record values in Chapter 5.

Sometimes in statistical analysis there is difference between registered distribution and the original distribution. Many events in nature are conveyed via the study of relation between registered and original distributions. Kullback Leibler [66] information measure is connected with the statistical problem of discrimination by considering a measure of distance or divergence between two distributions associated with the same random experiment. In Chapter 6, we have provided an extension of Kullback Leibler [66] information measure to k-record values. The distance between two k-record distributions of residual lifetime has found. Also keeping the record times fixed, we have derived the distance between k-record value and l-record value.

Over the time various generalizations and characterizations of Shannon entropy have been introduced by many researchers which are suitable for different types of problems. Various generalizations of Shannon entropy like Renyi, Havrda and Charvat are important in Probability and Statistics because of their role in large deviations theory and in the study of likelihood-based inference principles. Varma [122] entropy measure plays an essential role as a measure of complexity and uncertainty in different areas. We have extended the concept of Varma entropy to k-record values in Chapter 7.

Further Scope Of The Work

While assembling this thesis many ideas have emanated in our mind which can be useful for further study. The concept of weighted uncertainty measures based on record values is still untouched. A huge literature on weighted entropy measure and its generalizations is available, but no one yet has explored this idea using record values. We have tried to discuss the past entropy measure for k-record values but this idea for various uncertainty measures and its generalizations for the past lifetime of a random variable can be explored further.

Distribution function based uncertainty measures are of great importance. A huge literature is available in favour of this idea. Some work has also been carried using this idea for record values, but being a useful concept, it needs to be studied further in respect of record values.

In various realistic situations, the discrete random data is of great usage. But in comparison to continuous case, less work has been carried on discrete domain. We can study various information theoretic measures using discrete record random variables. The work presented in this thesis can be studied for bivariate and multivariate domains also.

The concept of k-records have been studied by many researchers, but still it can be explored further for various entropy measures.

Bibliography

[1] Abbasnejada, M. Arghamia, N. R. Morgenthalerb, S. and Borzadarana, G. R. M. (2010). On the dynamic survival entropy, Statistics and Probability Letters, **80** (23), 1962-1971.

[2] Abraham, B. and Sankaran, P. G. (2005). Renyi's entropy for residual life distribution, Statistical Papers, **46**, 17-30.

[3] Aczel, J. and Daroczy, Z. (1975). On Measures of Information and Their Characterizations, Academic Press, New York.

[4] Ahmadi J. and Arghami, N. R.(2003). Comparing the Fisher information in record values and iid observations, Statistics, **37**(5), 435-441.

[5] Ahmadi, J. and Balakrishnan, N. (2005). Preservation of some reliability properties by certain record statistics, A Journal of Theoretical and Applied Statistics, **39** (4), 347-354.

[6] Ahmadi J. and Doostparast, M. (2008). Statistical inference based on k-records, Iranian Journal of Numerical analysis and optimization, **1**(1),67-82.

[7] Ahmadi, M. and Mohtashami, B. (2013). Kullback-Leibler Information in View of an extended version of k-Records, **20**(1), 1-13.

[8] Ahsanullah, M. (2004). Record Values-Theory and Applications, University Press of America Inc., New York.

[9] Arnold, B. C. Balakrishnan, N. and Nagaraja, H. N. (1998). Records, Wiley, New York.

[10] Arghami, N. R. and Abbasnejad, M. (2011). Renyi entropy properties of order statistics, Communications in Statistics-Theory and Methods, **40**, 40-52.

[11] Asadi, M. and Ebrahimi, N. (2000). Residual entropy and its characterizations in terms of hazard function and mean residual life function, Statistics and Probability Letters, **49**, 263-269.

[12] Asadi, M. Ebrahimi, N. and Soofi, E. (2005). Dynamic generalized information measures, Statistics and Probability Letters, **71**, 89-98.

[13] Asadi, M. and Zohrevand, Y. (2007). On the dynamic cumulative residual entropy, Journal of Statistical Planning and Inference, **137** (6), 1931-1941.

[14] Asha, P. S. and Chacko, M. (2015). Residual renyi entropy of k-record values, Communications in Statistics- Theory and Methods, **45**(16), 4874-4885.

[15] Baig, M. A. K. and Dar, J. G. (2008). Generalized residual entropy function and its applications, European Journal of Pure and Applied Mathematics, **4**, 30-40.

[16] Balakrishnan, N. and Chan, P.S. (1998). On the normal record values and associated inference, Statistics and Probability Letters, **39**, 73-80.

[17] Balakrishnan, N. and Stepanov, A. (2004). Two characterizations based on order statistics and records, Journal of Statistical Planning and Inference, **124**, 273-287.

[18] Barlow, R. E. and Proschan, F. (1975). Statistical Theory of Reliability and Life Testing, Holt, Rinehart and Winston, New York.

[19] Baratpour, S. Ahmadi, J. and Arghami, N. R. (2007). Some characterizations based on entropy of order statistics and record values, Communications in Statistics-Theory and Methods, **36**, 47-57.

[20] Baratpour, S. Ahmadi, J. and Arghami, N. R. (2008). Characterizations based on Renyi entropy of order statistics and record values, Journal of Statistical Planning and Inference, **138**, 2544-2551.

[21] Belzunce, F. Navarro, J. Ruiz, J. M. and Aguila, Y. (2004). Some results on residual entropy function, Metrika , **59**, 147-161.

[22] Berred, M.(1995). k-record values and the extreme-value index, Journal of Statistical Planning and Inference, **45**, 49 - 63.

[23] Block, H. W. Savits, T. H. and Singh, H. (1998). The reversed hazard rate function, Probability in the Engineering and Informational Sciences, **12**, 69-90.

[24] Chandler, K. N.(1952). The distribution and frequency of record values, Journal of Royal Statistical society, **14**B, 220-228.

[25] Chandra, N. K. and Roy, D. (2001). Some results on reversed hazard rate, Probability in the Engineering and Informational Sciences, **15**, 95-102.

[26] Chandra, N. K. and Roy, D. (2005). Classification of distributions based on reversed hazard rate. Calcutta Statistical Association Bulletin, **56**, 221-224.

[27] Cover, T. M. and Thomas, J. (1991). The Elements of Information Theory, John Wiley and Sons.

[28] Cox, D. R. (1972). Regression models and life tables, Journal of the Royal Statistical Society, Series B, **34**, 187-220.

[29] David, H.A. and Nagaraja, H.N.(2003). Order Statistics, Wiley, New York.

[30] Di Crescenzo, A. (2000). Some results on the proportional reversed hazard model, Statistics and Probability Letters, **50**, 313-321.

[31] Di Crescenzo, A. and Longobardi, M. (2002). Entropy-based measure of uncertainty in past lifetime distributions, Journal of Applied Probability, **39**, 434-440.

[32] Di Crescenzo, A. and Longobardi, M. (2004). A measure of discrimination between past lifetime distributions, Statistics and Probability Letters, **67**, 173-182.

[33] Di Crescenzo, A. and Longobardi, M. (2006). On weighted residual and past entropies, Scientiae Mathematicae Japonicae, **64** (3), 255-266.

[34] Di Crescenzo, A. and Longobardi, M. (2009). On cumulative entropies, Journal of Statistical Planning and Inference, **139** (12), 4072-4087.

[35] Dziubdziela, W. and Kopocinski, B. (1976). Limiting properties of k-th record values, Zastosowania Matematyki, **15**, 187-190.

[36] Ebrahimi, N . (1996). How to measure uncertainty in the residual life time distribution, Sankhya A, **58**, 48-56.

[37] Ebrahimi, N. (1997). Testing whether lifetime distribution is decreasing uncertainty, Journal of Statistical Planning and Inference, **64**, 9-19.

[38] Ebrahimi, N. (2000). The maximum entropy method for lifetime distributions, Sankhya A, **62**, 236-243.

[39] Ebrahimi, N. and Kirmani, S.N.U.A. (1996 a). A characterization of the proportional hazards model through a measure of discrimination between two residual life distributions, Biomertika, **83** (1), 233-235.

[40] Ebrahimi, N. and Kirmani, S.N.U.A. (1996 b). A measure of discrimination between two residual life time distributions and its applications, The Annals of Mathematical Statistics, **48**, 257-265.

[41] Ebrahimi, N. and Pellerey, F. (1995). New partial ordering of survival functions based on the notion of uncertainty, Journal of Applied Probability, **32**, 202-211.

[42] Ebrahimi, N. Soofi, E. S. and Zahedi, H. (2004). Information properties of order statistics and spacings, IEEE Transactions of Information Theory, **50**, 177-183.

[43] Fashandi, M. and Ahmadi, J. (2012). Characterizations of symmatric distributions based on Renyi entropy, Statistics and Probability Letters, **82**, 798-804.

[44] Feller, W. (1966). An introduction to Probability theory and its applications, New York, NY. John Wiley and Sons.

[45] Finkelstein, M. S. (2002). On the reversed hazard rate, Reliability Engineering and System Safety, **78**, 71-75.

[46] Goel, R. Taneja, H.C. and Kumar, V. (2018). Kerridge measure of inaccuracy for record statistics, Journal of Information and Optimization Sciences, **39**(5), 1149-1161.

[47] Goel, R. Taneja, H.C. and Kumar, V. (2018). Measure of inaccuracy and k-record statistics, Bulletin of Calcutta Mathematical Society, **110** (2), 151-166.

[48] Goffman, C. and Pedrick, G. (1965). First course in functional analysis, Prentice Hall Inc., London.

[49] Gupta, R. C. and Gupta, R. D. (1998). Proportional reversed hazard rate model and its applications, Journal of Statistical Planning and Inference, **137** (11), 3525-3536.

[50] Gupta, R. D. and Kundu, D. (1999). Generalized exponential distributions, Australia and New Zealand Journal of Statistics, **41**(2), 173-188.

[51] Gupta, R. C. and Kirmani, S. N. U. A. (2008). Characterizations based on conditional mean function, Journal of Statistical Planning and Inference, **138**, 964-970.

[52] Gupta, R. C. Gupta, P. L. and Gupta, R. D. (1998). Modeling failure time data with lehman alternative, Communications in Statistics-Theory and methods, **27**, 887-904.

[53] Gupta, R. D. and Nanda, A. K. (2001). Some results on reversed hazard rate ordering, Communications in Statistics -Theory and Methods, **30**, 2447-2457.

[54] Gupta, R. C. and Wu, H. (2001). Analyzing survival data by proportional reversed hazard model, International Journal of Reliability and Applications, **2** (1), 1-26.

[55] Hartly, R. T. V. (1928). Transmission of information, Bell System Technical Journal, **7**, 535-563.

[56] Havrda, J. and Charvat, F. (1967). Quantification method of classification process: concept of structural $\alpha-$entropy, Kybernetika, **3**, 30-35.

[57] Hooda, D. S and Saxena, S. (2011). Generalized measures of discrimination between past life time distributions, Pakistan Journal of Statistics and Operation Research, **2**, 233-243.

[58] Johnson, D. H. and Glantz, R. M. (2004). When does interval coding occur?, Neurocomputing, **58** (60), 13-18.

[59] Kamps, U. (1995). A concept of generalized order statistics. B. G. Teubner, Stuttgart.

[60] Kapur, J. N. (1968). On information of order α and β, Proceeding of the Indian Academic Sciences, **48** (A), 65-75.

[61] Kapur, J. N. (1994). Measure of Information and Their Applications, Wiley Eastern Ltd., New Delhi.

[62] Kayid, M. and Ahmad, I. A. (2004). On the mean inactivity time ordering with reliability applications, Probability in the Engineering and Informational Sciences, **18** (3), 395-409.

[63] Kerridge, D. F. (1961). Inaccuracy and inference, Journal of the Royal Statistical Society, Series B, **23**, 184-194.

[64] Keilson, J. and Sumita, V. (1982). Uniform stochastic ordering and related inequalities, The Canadian Journal of Statistics, **10** (3), 181-191.

[65] Khinchin, A. J. (1957). Mathematical Foundation of Information Theory, Dover, New York .

[66] Kullback, S. and Leibler, R. A. (1951). On information and sufficiency, Annals of Mathematics Statistics, **22**, 79-86.

[67] Kullback, S. (1959). Information Theory and Statistics, Wiley, New York.

[68] Kumar, V. Taneja, H. C. and Srivastava, R. (2010). A dynamic measure of inaccuracy between two past lifetime distributions, Metrika, **74** (1), 1-10.

[69] Kumar, V. Taneja, H. C. and Srivastava, R. (2010). Non-additive entropy measure based residual lifetime distributions, JMI International Journal of Mathematical Sciences, **1**(2), 1-9.

[70] Kumar, V. Taneja, H. C. and Srivastava, R. (2011). On dynamic Renyi cumulative residual entropy measure, Journal of Statistical Theory and Applications, **10**(3), 491-500.

[71] Kumar, V. (2015). Generalized entropy measure in record values and its applications, Statistics and Probability Letters, **106**, 46-51.

[72] Kumar, V. and Taneja, H. C. (2015). Non-additive entropy measure and record values, Applied Mathematics and Information Sciences, **9** (3), 1541-1548.

[73] Kundu, C. Nanda, A. K. and Maiti, S. S. (2010). Some distributional results through past entropy, Journal of Statistical Planning and Inference, **140** (5), 1280-1291.

[74] Lehmann, E. L. (1953). The power of rank test, Annals of Mathematical Statistics, **24**, 23-43.

[75] Leemis, L. M. and Beneke, M. (1990). Burn-In models and methods: A Review, IIE Transactions, **22**, 172-180.

[76] Madadi, M. and Tata, M. (2011). Shannon information in record data, Metrika, **74**, 11-31.

[77] Madadi, M. and Tata, M. (2014). Shannon Information In k-records, Communications in Statistics - Theory and Methods, **43**(15), 3286-3301.

[78] Malinowska, I. and Szynal, D. (2004). On a family of Bayesian estimators and predictors for Gumbel model based on the k-th lower record, Applicationes Mathematicae, **31**(1), 107- 115.

[79] Mary, M. S. and Chacko, M. (2010). Estimation of uniform distribution based on k-record values, Calcutta Statistical Association Bulletin, **62**, 143-158.

[80] McEliece, R. J. (1977). The theory of information and coding, A Mathematical Framework For Communication. Addison-Wesley, Reading, Mass.

[81] Mosayeb, M. and Borzadaran G.R., M. (2013). Kullback-Leibler information in view of an extended version of k-records, Communications for Statistical Applications and Methods, **20**(1), 1-13.

[82] Muth, E. J. (1977). The null distribution for a test of constant versus bathtub failure rate, Scandinavian Journal of Statistics, **12**, 55-61.

[83] Nath, P. (1968). Inaccuracy and coding theory, Metrika, **13**, 123-135.

[84] Nair, N. U. and Gupta, R. P. (2007). Characterization of proportional hazard models by properties of information measures, International Journal of Statistical Sciences, **6**, 223-231.

[85] Nair, N. U. Sankaran, P. G. and Asha, G. (2005). Characterizations of distributions using reliability concepts, Journal of Applied Statistical Science, **14**, 237-242.

[86] Nanda, A. K. and Das, S. (2011). Dynamic proportional hazard rate and reversed hazard rate models, Journal of Statistical Planning and Inference, **141** (6), 2120-2127.

[87] Nanda, A. K. and Paul, P. (2006). Some properties of past entropy and their applications, Metrika, **64** (1), 47-61.

[88] Nanda, A. K. and Paul, P. (2006). Some results on generalized residual entropy, Information Sciences, **176**, 24-47.

[89] Navarro, J. Sunoj, S. M. and Linu, M. N. (2011). Characterizations of bivariate models using dynamic Kullback-Leibler discrimination measures, Statistics and Probability Letters, **81**, 1594-1598.

[90] Nyquist, H. (1924). Certain factors affecting telegraph speed, Bell System Technical Journal, **3**, 324-346.

[91] Nyquist, H. (1928). Certain topics in telegraph transmission theory, Transactions of the American Institute of Electrical Engineers, **47**(2), 617-644.

[92] Park, S. (1995). The entropy of consecutive order statistics, IEEE Transactions on Information Theory, **41**(6), 2003-2007.

[93] Psarrakos, G. and Navarro, J. (2013). Generalized cumulative residual entropy and record values, Metrika, **76** (5), 623-640.

[94] Rao, M. (2005). More on a new concept of entropy and information, Journal of Theoretical Probability, **18**, 967-981.

[95] Rao, M. Chen, Y. Vemuri, B. C. and Wang, F. (2004). Cumulative residual entropy: a new measure of information, IEEE Transactions on Information Theory, **50**, 1220-1228.

[96] Rajesh, G. and Nair, K. R. M. (1998). Residual entropy function in discrete time, Far East Journal of Theoretical Statistic, **2** (I), 1-10.

[97] Razmkhah, M. Morabbi, H. and Ahmadi, J. (2012). Comparing two sampling schemes based on entropy of record statistics, Statistical Papers, **53**, 95-106.

[98] Raqab, M. Z. and Awad, A. M. (2000). Characterizations of the Pareto and related distributions, Metrika, **52**, 63-67.

[99] Renyi, A. (1961). On measures of entropy and information, Proceeding of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, university of California Press, Berkeley, 547-561.

[100] Sankaran, P. G. and Gupta, R. P. (1999). Characterization of lifetime distributions using measure of uncertainty, Calcutta Statistical Association Bullettin, **49**(3), 159-166.

[101] Schneider, T. D. (1997). Information content of individual genetic sequences, Journal of Theoretical Biology, **189**(4), 427-441.

[102] Sengupta, D. Singh, H. and Nanda, A. K. (1999). The proportional reversed hazards model, Technical Report, Applied Statistics Division, Indian statistical Institute, Kolkata.

[103] Shaked, M. and Shanthikumar, J.G. (2007). Stochastic Orders, Springer, New York.

[104] Shannon, C. E. (1948). A mathematical theory of communication, Bell System Technical Journal, **27**, 379-423 and 623-656.

[105] Sharma, B. D. and Taneja, I. J. (1975). Entropy of type (α, β) and other generalized measures in information theory, Metrika, **22**, 205-215.

[106] Su, J.C. Su, N.C. and Huang, W.J. (2008). Characterizations based on record values and order statistics, Journal of Statistical Planning and Inference, **138**, 1358-1367.

[107] Sultan, K. S. Moshref, M. E. and Childs, A. (2002). Record values from generalized power function distribution and associated inference, Journal of Applied Statistical Science, **11**,143-156.

[108] Sullo, P. and Rutherford, D. K. (1977). Characterizations of the power distribution by conditional excedence, The American Statistical Association Proceedings of the Business and Economics, 787-792.

[109] Sunoj, S. M. and Linu, M. N. (2012). Dynamic cumulative residual Renyi's entropy, Statistics, **46** (1), 41-56.

[110] Sunoj, S. M. Sankaran, P. G. and Maya, S. S. (2009). Characterizations of life distributions using conditional expectations of doubly (interval) truncated random variables, Communications in Statistics-Theory and Methods, **38**, 1441-1452.

[111] Swartz, G. B. (1973). The mean residual life time function, IEEE Transactions on Reliability, **22**, 108-109.

[112] Tahmasebi, S. and Eskandarzadeh, M. (2017). Genaeralized cumulative entropy based on kth lower record values, Statistics and Probability Letters, **126**, 164-172.

[113] Taneja, H.C. (1985). On measure of relative useful information, Kybernetika, **21**, 148-156.

[114] Taneja, I. J. (1995). New developments in generalized information measures, Advances in Imaging and Electron Physics, **91**, 37-136.

[115] Taneja, I. J. (1989). On generalized information measures and their applications, Advances in Electronics and Electron Physics, **76**, 327-413.

[116] Taneja, H.C. and Kumar, V. (2012). On dynamic cumulative residual inaccuracy measure, Proceedings at World Congress of Engineering, WCE-Vol. I, held at London, U.K., July 06-08.

[117] Taneja, H. C. Kumar, V. and Srivastava, R. (2009). A dynamic measure of inaccuracy between two residual lifetime distributions, International Mathematical Forum, **4** (25), 1213-1220.

[118] Taneja, H. C. (1985). On measure of relative useful information, Kybernetika, **21** (2), 148-156.

[119] Tsallis, C. and Brigatti, E. (2004). Non-extensive statistical mechanics: a brief introduction, Continuum Mechanics and Thermodynamics, **16**, 223-235.

[120] Thapliyal, R. and Taneja, H. C. (2015). On Renyi entropy of order statistics, International Journal of Biomathematics, **8**(6), 1550080(1-10).

[121] Thapliyal, R. Taneja, H. C. and Kumar, V. (2015). Characterization results based on non-additive entropy of order statistics, Physica A: Statistical Mechanics and Its Applications, **417**, 297-303.

[122] Varma, R. S. (1966). Generalizations of Renyi's entropy of order α, Journal of Mathematical Sciences, **1**, 34-48.

[123] Wells (2009). Applied Coding and Information Theory for Engineers, Pearson Education.

[124] Wong, K. M. and Chen, S. (1990). The entropy of ordered sequences and order statistics, IEEE Transactions of Information Theory, **36**, 276-284.

[125] Xekalaki, E. (1983). Hazard function and life distributions in discrete time, Communications in Statistics-Theory and Methods, **12**, 2503-2509.

[126] Zahedi, H. and Shakil, M. (2006). Properties of entropies of record values in reliability and life testing context, Communications in Statistics - Theory and Methods, **35**, 997-1010.

[127] Zarezadeh, S. and Asadi, M. (2010). Results on residual Renyi entropy of order statistics and record values. Information Sciences, **180**, 4195-4206.

List Of Publications

1. **Ritu Goel**, H.C. Taneja and Vikas Kumar. *Measure of Entropy for Past Lifetime and k-Record Statistics*, *Physica A*, 503, 623-631, (2018).

2. **Ritu Goel**, H.C. Taneja and Vikas Kumar. *Measure of Inaccuracy and k-Record Statistics*, *Bulletin of Calcutta Mathematical Society*, 110 (2), 151-166, (2018).

3. **Ritu Goel**, H.C. Taneja and Vikas Kumar. *Kerridge Measure of Inaccuracy for Record Statistics*, *Journal of Information and Optimization Sciences*, 39(5), 1149-1161, (2018).

4. **Ritu Goel**, H.C. Taneja and Vikas Kumar. *A Measure of Discrimination Between Two Residual Lifetime Distributions For k-Record Values*, Communicated.

5. **Ritu Goel**, H.C. Taneja and Vikas Kumar. *Cumulative Residual Inaccuracy Measure for k-Record Values*, Communicated.

6. Vikas Kumar, H.C. Taneja and **Ritu Goel**. *On Generalized Information Measure of Order (α, β) and k-Record Statistics*, Communicated.

Papers Presented in Conferences

1. **Ritu Goel**, H.C.Taneja and Vikas Kumar. *Kerridge measure of inaccuracy for record statistics*, Presented at International Conference on Interdisciplinary Mathematics, Statistics and Computational Techniques held at Manipal University, Jaipur, Dec 22-24, 2016.

2. **Ritu Goel**, H.C.Taneja and Vikas Kumar. *Measure of inaccuracy and k-record statistics*, Presented at National Seminar on Recent Developments in Mathematical Sciences held at MDU, Rohtak, Mar 07-08, 2017.

3. **Ritu Goel** and H.C.Taneja. *Measure of Discrimination between Two Residual Lifetime Distributions for k-Record Values*, Presented at International Research Symposium on Engineering and Technology held at Singapore, Aug 28-30, 2018.

4. **Ritu Goel** and H.C.Taneja. *On Cumulative Inaccuracy*, Presented at International Conference Recent Advances in Pure and Applied Mathematics held at Delhi Technological University, Delhi, Oct 23-25, 2018.