

Chapter 1

INTRODUCTION

The Shannon’s entropy, introduced by Shannon (1948), has been extensively used as a quantitative measure of uncertainty associated with a random phenomena. If $A_1, A_2... A_n$ are mutually exclusive and exhaustive events in a sample space with respective probabilities $p_1, p_2... p_n$, the Shannon’s entropy is defined as

$$H_n (P) = - \sum_{i=1}^{n} p_i \log p_i.$$ 

$H_n (P)$ is being interpreted as a measure of uncertainty concerning the outcome of the experiment or a measure of information conveyed through the knowledge of the probabilities associated with the events.

Observing that the Shannon’s entropy satisfies several properties, the earlier work on the Shannon’s entropy was centered around characterizing $H_n (P)$ based on several postulates. The works of Khinchin (1953), Tverberg (1958), Chaundy and Mcleod (1960), Lee (1964), Mathai and Rathie (1975), Ebanks et al. (1998), Yeung (2002), and Csiszar (2008) in this direction. Another aspect of interest that has received much attention among researchers is the identification of probability distributions that maximizes the Shannon’s entropy subject to some restrictions on the underlying random variable. The books by Kapur (1989, 1994) provide a more or less exhaustive review of various maximum entropy models.

In the continuous setup if $f(x)$ denotes the probability density function of a random variable $X$ with support $[a,b]$, the continuous analogue of Shannon’s entropy takes the form
Introduction

\[ H(F) = -\int_a^b f(x) \log f(x) \, dx. \]

Ebrahimi and Pellerey (1995) have extended the definition of Shannon’s entropy to the left truncated situation and they used this measure to introduce a new partial ordering for life distributions. Ebrahimi (1996) has given an upper bound for this measure in terms of the mean residual life function, \( m(t) \), namely

\[ H(F; t) \leq 1 + \log m(t), \]

where \( m(t) < \infty \). Nair and Rajesh (1998), Sankaran and Gupta (1999), Asadi and Ebrahimi (2000) and Belzuence et al. (2004) have looked into the problem of characterizing probability distributions using the functional form of the residual entropy function. Rajesh and Nair (1998) have defined the residual entropy function in discrete time domain and have shown that it determines the distribution uniquely. Further, it is established that the constancy of the same is characteristic to the geometric distribution. Di Crescenzo and Longobardi (2002) have shown that in many realistic situations uncertainty is not necessarily related to the future but can also refer to the past. For considering such situations, they proposed the past entropy defined over \( (0, t) \). Recently, Nanda and Paul (2006,a) have proposed some ordering properties based on this measure.

Kullback and Leibler (1951) have extensively studied the concept of directed divergence, which aims at discrimination between two populations. Aczel and Daroczy (1975) laid down an axiomatic foundation to this concept. Ebrahimi and Kirmani (1996, b) extended this concept and has given a measure of discrimination between two residual lifetime distributions. Further, they proved that the constancy of this measure is a characteristic property of the proportional hazards model. Along the similar lines of the measure proposed by Ebrahimi and Kirmani (1996, b), Di Crescenzo and Longobardi (2004) have examined the problem of discrimination between the past lives.
Another useful measure for discrimination among distributions is the notion of affinity studied by Matusita (1954). Affinity focuses attention on the likeness of distribution and has properties similar to that of Kullback-Leibler divergence measure. Kirmani (1968) has shown that the affinity between two distributions is related to the idea of distance between distributions. In testing hypothesis, it is desirable to know bounds of errors, because even if the most powerful test is adopted, it is often the case that we cannot obtain the exact value of the power of the test. However, we can easily get them in terms of affinity. The relative Renyi entropy, also known as Chernoff distance, finds application in several branches of learning as a potential measure of distance between two populations. Asadi et. al (2005) have studied the application of this measure in the context of reliability studies.

The notion of inaccuracy was introduced by Kerridge (1961) and can be viewed as a generalization of the Shannon’s entropy. Suppose that the experimenter asserts that the probability of the $i^{th}$ eventuality is $q_i$ when the true probability is $p_i$. Then the inaccuracy of the observer, as proposed by Kerridge (1961), can be measured by

$$I(P, Q) = -\sum_{i=1}^{n} p_i \log q_i,$$

where $P = (p_1, p_2, ..., p_n)$ and $Q = (q_1, q_2, ..., q_n)$ are two discrete probability distributions such that $p_i \geq 0$, $q_i \geq 0$ and $\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i = 1$. In fact, the Kerridge’s inaccuracy measure can be expressed as the sum of a measure of uncertainty and a measure of discrimination between two populations. When an experimenter states the probabilities of various events in an experiment, the statement can lack precision in two ways: one is resulting from incorrect information and the other from vagueness in the statement. Kerridge (1961) proposed the “inaccuracy measure” that can take accounts for these two types of errors. Nath (1968) extended Kerridge’s inaccuracy to the case of continuous situation and discussed some properties. If $F(x)$ is the actual
distribution function corresponding to the observations and \( G(x) \) is the distribution function assigned by the experimenter and \( f(x) \) and \( g(x) \) are the corresponding density functions, the inaccuracy measure is defined as

\[
I(F, G) = -\int_0^\infty f(x) \log g(x) dx.
\]

He also extended this measure to inaccuracies of order ‘\( r \)’. Nair and Gupta (2007) extended the definition of measure of inaccuracy to the truncated situation. Recently, Taneja et. al (2009) proposed the uniqueness property of the dynamic inaccuracy measure defined by Nair and Gupta (2007) and some properties of this measure were also studied. In addition, the concept of inaccuracy has its application in statistical inference, estimation and coding theory.

Even though concepts such as failure rate, mean residual life function, vitality function etc are extensively used in reliability studies for modeling lifetime data, recently a lot of interest has evoked in using entropy concepts to describe the stability of components. In life time studies, the data is generally truncated. Hence there is scope for extending information theoretic concepts to the truncated situation. Motivated by this, in the present study, we extend the definition of inaccuracy, affinity and Chernoff distance to the truncated situation. Further we also look into the problem of characterization of probability distributions using the functional form of these measures.

After the present introductory chapter, in Chapter 2 we give a brief review of the existing literature in the area of study. In Chapter 3, we extend the definition of inaccuracy to the truncated situation and provide characterization results for certain probability distributions. The inaccuracy measure is generalized to inaccuracies of order ‘\( r \)’ in Chapter 4. Characterizations of distributions in the context of proportional hazards model and proportional reversed hazards model using the functional form of the generalized inaccuracy measure are also given in this chapter. In Chapter 5, we extend the notion of Chernoff distance to the truncated situation and obtain characterization results using functional form of the
truncated Chernoff distance. We also discuss affinity in the truncated situation, which is a special case of the Chernoff distance, in this chapter. Residual inaccuracy measure and affinity in discrete setup are the subject matter of Chapter 6. Towards the end of this chapter we also give a plan for future study in this area.