INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

ENTROPY FOR DIFFERENT FRAILTY DISTRIBUTIONS

ARVIND PANDEY, MAHESH GODARA

Department of Statistics, Central University of Rajasthan, Ajmer-305817, India.

Accepted Date: 30/10/2017; Published Date: 01/11/2017

Abstract: - Entropy is the measure of disorder in physical systems, or an amount of information that may be gained by observations of disordered systems. Entropy, as a measure of randomness contained in a probability distribution, is a fundamental concept in information theory. Entropy is simply the average (expected) amount of the information from the event. The entropy measures the uniformity of a distribution. The concept of frailty provides a suitable way to introduce random effects in the model to account for association and an unobserved heterogeneity. In its simplest form, a frailty is an unobserved random factor that modifies multiplicatively the hazard function of an individual or a group or cluster of individuals. We find the entropies of different frailty models.

Keywords: Frailty Models, Shananos Entropy, Residual Entropy and Past Entropy.

Corresponding Author: MR. ARVIND PANDEY

Access Online On:

www.ijpret.com

How to Cite This Article:

Arvind Pandey, IJPRET, 2017; Volume 6 (3): 19-42

PAPER-QR CODE

INTRODUCTION

Entropy is the measure of disorder in physical systems, or an amount of information that may be gained by observations of disordered systems. Entropy, as a measure of randomness contained in a probability distribution, is a fundamental concept in information theory. In the recent past, many researchers have taken a keen interest in the measurement of uncertainty associated with a probability distribution of particular interest in probability and statistics is the notion of entropy, introduced by [1]. Entropy is simply the average (expected) amount of the information from the event. Shannon entropy, we had this wonderful intuition in which it represented the `information content' of a discrete distribution. That is, Shannon entropy could also be defined as "the expected value of the information of the distribution" or the number of bits you'd need to reliably encode n symbols. In the continuous case, this intuition deteriorates as H(X) does not give you the amount of information in X. Generaly entropy denoted by H(f).The entropy measures the uniformity of a distribution. As H(f) increases, f(x) approaches to uniform. Consequently, the concentration of probabilities decreases and it becomes more difficult to predict an outcome of a draw from f(x). In fact, a very sharply peaked distribution has a very low entropy, whereas if the probability is spread out the entropy is much higher. In this sense H(X) is a measure of uncertainty associated with f(x). Shannon was the first to introduce entropy, known as Shannons entropy or Shannons information measure, into information theory. The properties and virtues of entropy have been thoroughly investigated by [1] and [2]. Further, many generalizations of entropy have been proposed by [3], [4], [6], [7], [8] and [9]. In recent years, the modification of Shannon entropy as a measure of uncertainty in residual lifetime distributions has drawn attention of many researchers (cf. [10]; [11]. [12] have introduced past entropy over (0, t), since it is reasonable to presume that in many realistic situations uncertainty is not necessarily related to the future but can also refer to the past. They have also shown the necessity of past entropy and its relation with the residual entropy.

Entropy measures the extent to which observations are concentrated around a single point, and is a descriptive measure similar to, but not the same as, standard deviation. For the multinomial distribution, entropy is referred to as the heterogeneity present in the variable. The concept of entropy is illustrated in Figure 1 using two extreme cases. Suppose we have a discrete variable X that can take one of three possible values. Figure 1 (a) shows the probability of an observation belonging to the second category, and since this probability is 1, there is no uncertainty as to where the next observed value will be Maximum certainty is attained, which is also referred to as minimum uncertainty or complete homogeneity. Figure 1 (b) shows that the next observed value is equally likely to belong to any of the categories. There is no certainty, also called minimum certainty, maximum uncertainty or complete heterogeneity.

2. Frailty

The statistical analysis of survival data is an important topic in many areas, including medicine, epidemiology, biology, demography, economics, engineering and other fields. A variety of techniques have been developed to analysis survival data. A common approach to the analysis of survival data is based on the assumption that the study population is homogeneous. That is, conditional on the covariates, every individual has the same risk of experiencing an event such as death or disease recurrence. Proportional hazards models and accelerated failure time models are classical models that are frequently used for the analysis of univariate (censored) survival data. Standard survival models assume independence between the survival times. Frailty models provide a useful extension of the standard survival models by introducing a random effect (frailty) when the survival data are correlated. Frailty models can be used in the survival analysis to represent random effects or unexplained heterogeneity between individuals or groups. Multivariate or cluster failure time data are commonly encountered in the survival analysis, and finding an appropriate method to model the correlation among the observations is a very important issue. Frailty model provide an appropriate method to model the correlation among the multivariate data. Multivariate data occurs for example if lifetimes (or times of onset of a disease) of relatives (twins, parent-child) or recurrent events like infections in the same individual are considered. In such cases independence between the clustered survival times cannot be assumed. A convenient choice for modeling the correlation in survival data is the frailty model, which was originally proposed by [13]. The term frailty was first introduced by [13] in univariate survival models. Frailty models extend the Proportional Hazard model by including random effects, called frailty, to account for dependency between observations. However, it is not always reasonable to assume that all the individuals in the same cluster share exactly the same frailty. An extension of the shared frailty model is the correlated frailty model where individuals in the same cluster have different, yet correlated, frailties.

2.1 Frailty Models

The concept of frailty provides a suitable way to introduce random effects in the model to account for association and an unobserved heterogeneity. In its simplest form, a frailty is an unobserved random factor that modifies multiplicatively the hazard function of an individual or a group or cluster of individuals.

2.2 Consequences of Ignoring Frailty

It is very important to consider the effect of ignoring frailty where the existence of heterogeneity may be present. The impact of frailty in event history models differs substantially from the impact of frailty in linear regression models. In ordinary regression models, unobserved heterogeneity leads to more variability of the response compared to the case when the variables are included. In event history data, however, the increased variability implies a change in the hazard function. When the hazard rate exhibits positive duration dependence, ignoring frailty will make this duration dependence, less pronounced or even negative. When the hazard rate exhibits negative duration dependence, ignoring frailty will make this negative duration dependence stronger. Another consequence of ignoring frailty is that the effect of a covariate is biased towards zero. The consequence of ignoring frailty factors, which is very important in modeling stochastic risk factor, is that intensity estimation in credit risk models would be biased. Therefore, the exposure at the default cannot be integrated into the credit portfolio model. This is especially a problem for market-driven instruments, such as interest rate derivatives. It is recognised in the field of econometrics and biometrics, through empirical evidence, that if frailty is present but ignored then covariate effects will be underestimated ([14]; [15]). [14] confirmed this evidence for uncensored Weibull survival data through theoretical work. [?] showed that fitting misspecified Cox proportional hazards models to the marginal distributions (ignoring frailty) leads to regression coefficient estimates biased towards zero by an amount which depends on the variability of the frailty terms and the form of frailty distribution. He also concluded that the fitted marginal survival curves can also differ substantially from the true marginals. In the analysis of multivariate failure time data, failure to account for dependency has been shown to lead to biased parameter estimators. Moreover, ignoring frailty effects with finite mean may result in a negative bias in the estimated time dependence. Epidemiological relative risk measures are vulnerable to the frailty phenomenon, as frailty reduces the inuence of known covariates on the relative risk. Consequently, this can lead to non-proportional hazards.

3. Modeling Frailty

The generalization of the Cox proportional hazards model is the best and widely applied model that allows for the random effect by multiplicatively adjusting the baseline hazard function. Frailty models extend Cox proportional hazards model by introducing unobserved frailties to the model. In this case, the hazard rate will not be just a function of covariates, but also a function of frailties. A frailty model is a random effects model which has a multiplicative effect on the hazard rates of all the members of the subgroups. In an univariate survival models, it can be used to model the heterogeneity among individuals, which is the inuence of an unobserved risk factors in a proportional hazards model. In multivariate survival models, shared frailty model is used to model the dependence between the individuals in the group. In the multivariate case, unobserved frailty is common to a group of individuals.

In a univariate frailty model, let a continuous random variable T be a lifetime of an individual and the random variable Z be frailty variable. The conditional hazard function for a given frailty variable, Z = z at time t > 0 is,

$$h(t|z) = zh_0(t)e^{X\beta} \qquad (3.1)$$

where $h_0(t)$ is a baseline hazard function at time t > 0, X is a row vector of covariates,

and β is a column vector of regression coefficients. The conditional survival function for given frailty at time t > 0 is,

$$S(t) = \int_0^\infty S(t|z)f(z)dz$$

$$= \int_0^\infty e^{-zH_0(t)e^{X\beta}} f(z)dz$$

$$= L_z(H_0(t)e^{X\beta}) \qquad (3.3)$$

where $L_Z(.)$ is the Laplace transformation of the distribution of Z. Once we get the

function at time t > 0, of life time random variable for an individual, we can obtain probability structure and make their inferences based on it.

3.1 Gamma Frailty Model

Let a continuous random variable Z follows a gamma distribution. For identi_ability,

we assume Z has expected value equal to one which introduces restriction on the scale and the shape parameters. Under the restriction, the density function and the Laplace transformation of a gamma distribution reduces to,

$$f_z(z) = \begin{cases} \dfrac{\left(\frac{1}{\theta}\right)^{\frac{1}{\theta}}}{\Gamma\left(\frac{1}{\theta}\right)} z^{\frac{1}{\theta}-1} e^{-\frac{x}{\theta}} & ; \ z > 0, \theta > 0 \\ 0 & \end{cases} \qquad (3.4)$$

and $L_s(s) = (1+\theta s)^{-\frac{1}{\theta}}$ with variance of Z as θ.

3.2 Shannon's entropy

An intuitive understanding of information entropy relates to the amount of uncertainty about an event associated with a given probability distribution. As an example, consider a box containing many coloured balls. If the balls are all of different colours and no colour predominates, then our uncertainty about the colour of a randomly drawn ball is maximal. On the other hand, if the box contains more red balls than any other colour, then there is slightly less uncertainty about the result: the ball drawn from the box has more chances of being red (if we were forced to place a bet, we would bet on a red ball). Telling someone the colour of every new drawn ball provides them with more information in the first case than it does in the second case, because there is more uncertainty about what might happen in the first case than there is in the second. Intuitively, if we know the number of balls remaining, and they are all of one color, then there is no uncertainty about what the next ball drawn will be, and therefore there is no information content from drawing the ball. As a result, the entropy of the "signal" (the sequence of balls drawn, as calculated from the probability distribution) is higher in the first case than in the second.

→ Shannon, in fact, defined entropy as a measure of the average information content associated with a random outcome.

→ Shannon's definition of information entropy makes this intuitive distinction mathematically precise. His definition satisfies these desiderata:

→ The measure should be continuous i.e., changing the value of one of the probabilities by a very small amount should only change the entropy by a small amount.

→ If all the outcomes (ball colours in the example above) are equally likely, then entropy should be maximal. In this case, the entropy increases with the number of outcomes.

→ If the outcome is a certainty, then the entropy should be zero.

→ The amount of entropy should be the same independently of how the process is regarded as being divided into parts.

Shannon entropy is may be used globally, for the whole data, or locally, to evaluate entropy of probability density distributions around some points. This notion of entropy can be generalized to provide additional information about the importance of specific events, for example outliers or rare events. Comparing entropy of two distributions, corresponding for example to two features, Shannon entropy assumes implicite certain tradeoff between contributions from the tails and the main mass of this distribution. It should be worth- while to control this tradeoff explicitly, as in many cases it may be important to distinguish weak signal overlapping with much stronger one. Entropy measures that depend on powers of probability, $\sum_{i=1}^{n} P(x_i)^{\alpha}$ provide such control. If has large positive value this measure is more sensitive to events that occur often, while for large negative it is more sensitive to the events which happen seldom.

The Shannon's entropy associated with the random variable X is defined as

$$H(t) = E[-\log(f(t))]$$

Or

$$H(t) = -\int_0^{\infty} f(t)\ \log(f(t))dt$$

3.3 Renyi Entropy

The Rnyi entropy is named after Alfrd Rnyi. The Rnyi entropy is important in ecology and statistics as indices of diversity. The Rnyi entropy is also important in quantum information, where it can be used as a measure of entanglement. History: Alfred Renyi was looking for the most general definition of information measures that would preserve the additivity for independent events and was compatible with the axioms

of probability. He started with Cauchys functional equation: If p and q are independent than I(pq) = I(p) + I(q). Apart from a normalizing constant this is compatible with Hartleys information content I(p) = -log(p). If we assume that the events $X = x_1, x_{\ldots\ldots} x_N$ have different probabilities $p_1, p_2 \ldots\ldots p_N$ and each delivers Ik bits of information, then the total amount of information for the set is

$$I(p) = \sum_{k=n}^{N} P_k I_k$$

This can be recognized as Shannons entropy. But he reasoned that there is an implicit assumption used in this equation: we use the linear average, which is not the only one that can be used.

3.4 Special cases of the Rnyi entropy

As approaches zero, the Rnyi entropy increasingly weighs all possible events more equally, regardless of their probabilities. In the limit for 0, the Rnyi entropy is just the logarithm of the size of the support of X. The limit for 1 is the Shannon entropy. As approaches infinity, the Rnyi entropy is increasingly determined by the events of highest probability.

(1) Hartley or max-entropy:- Provided the probabilities are nonzero, H_0 is the logarithm of the cardinality of X, sometimes called the Hartley entropy of X,

$H_0(X) = \log n = \log|x|$.

(2) Shannon entropy:- The limiting value of H_α as 1 is the Shannon entropy.

$H_1(X) = -\sum_{i=1}^{n} p_i \log p_i$

(3) Collision entropy:- Collision entropy, sometimes just called "Rnyi entropy," refers to the case = 2,

$H_2(X) = -log \sum_{i=1}^{n} {p_i}^2 = -\log P(X = Y)$ where X and Y are independent and identically distributed.

(4) Min-entropy:-

Main article: Min entropy In the limit as $\alpha \to \infty$ the Rnyi entropy H_α converges to the min-entropy H_∞ :

$H_\infty(X) = \min(-\log p_i) = -(\max \log p_i) = -\log \max p_i$.

Equivalently, the min-entropy $H_\infty(X)$ is the largest real number b such that all events occur with probability at most 2^{-b}.

The name min-entropy stems from the fact that it is the smallest entropy measure in the family of Rnyi entropies. In this sense, it is the strongest way to measure the information content of a discrete random variable. In particular, the min-entropy is never larger than the Shannon entropy.

The min-entropy has important applications for randomness extractors in theoretical computer science: Extractors are able to extract randomness from random sources that have a large min-entropy; merely having a large Shannon entropy does not suffice for this task. if X is a random variable with density function f(x), the Renyi entropy is a measure of the uncertainty of the random variable. which is given by

$$H(p) = \frac{1}{1-p} \log \int_{-\infty}^{\infty} [f(x)]^p dx$$

Past entropy:-

Information theory includes the study of uncertainty measures plays a significant role in studying the various aspects of a system when it fails between two time points. In reliability theory and survival analysis, the residual entropy was considered in [11], which basically measures the expected uncertainty contained in remaining lifetime of a system. The residual entropy has been used to measure the wear and tear of components and to characterize, classify and order distributions of lifetimes by [16] and [10]. The notion of past entropy, which can be viewed as the entropy of the inactivity time of a system, was introduced in [13]. Many ageing (lifetime) distributions have been constructed with a view for applications in various disciplines, in particular, in reliability engineering, survival analysis, demography, actuarial study and others. Statistical analysis of lifetime data is an important topic in biomedical science, reliability engineering, social sciences and others. Typically, lifetime refers to human life length, the life span of a device before it fails, the survival time of a patient with serious disease from the date of diagnosis or major treatment or the duration of a social event such as marriage.

Let X be an absolutely continuous non- negative random variable having distribution function F(x) and survival function R(x) . The basic measure of uncertainty is defined by Shannon [1948] as

$$H(t) = -\int_0^{\infty} f(t)\log(f(t))\,dt$$

where f(x) is the density function of X . (Throughout this thesis, log will denote the natural logarithm). If X is a discrete random variable taking the values x_1, x_2, ……x_n with respective probabilities p_1,p_2,…..p_n then the Shannons entropy is defined as

$$H(p) = H(\mathrm{p}_1, p_2 \ldots . p_n) = -\sum_{k=1}^{n} p_k \log[p_k]$$

The role of differential entropy as a measure of uncertainty in residual life time distribution has attracted increasing attention in recent years. As urged by Ebrahimi [1996] , if a unit is known to have survived to age t, then H(X) is no longer useful for measuring the uncertainty about the remaining life time of the unit. Accordingly, he introduced the measure of uncertainty of residual life time distribution, H(X; t) of a component as,

H(t)=-$\int_0^{\infty} \frac{f(x)}{R(t)} \log(\frac{f(x)}{R(t)})dx$

Since this entropy is not applicable to a system that has survived for some units of time, the concept of residual entropy has been developed in the literature. [10] (1995) and [11]) introduced the concept of residual entropy in terms of conditional Shannon's measure. For a non-negative random variable X representing the life time of a component, the residual entropy function is the Shannon's entropy associated with the random variable X and is defined as

$$H(t) = -\int_t^{\infty} \frac{f(x)}{\bar{F}(t)} \log\left(\frac{f(x)}{\bar{F}(t)}\right) dx$$

Here f(x) is probability density function of distribution. $\bar{F}$ is survival function of distribution.

Cumulative Residual Entropy:-

We propose an alternative measure of uncertainty in a random variable X and call it the Cumulative Residual Entropy (CRE) of X .The main objective of our study is to extend Shannon entropy to random variables with continuous distributions. The concept we proposed overcomes the problems mentioned above, while retaining many of the important properties of Shannon entropy. For instance, both are decreased by conditioning, while increased by independent addition. They both obey the data processing inequality etc. However, the differential entropy doesn’t have the following important properties of CRE.

1) CRE has consistent definitions in both the continuous and discrete domains;

2) CRE is always non-negative;

3) CRE can be easily computed from sample data and these computations asymptotically converge to the true values.

4) The conditional CRE of X given Y is zero, if and only if X is a function of Y.

The basic idea in our definition is to replace the density function with the cumulative distribution in Shannons definition. The distribution function is more regular than the density function, because the density is computed as the derivative of the distribution. Moreover, in practice what is of interest and/or measurable is the distribution function. For example, if the random variable is the life span of a machine, then the event of interest is not whether the life span equals t, but rather whether the life span exceeds t. Our definition also preserves the well established principle that the logarithm of the probability of an event should represent the information content in the event. The discussions about the properties of CRE in next few sections, we trust, are convincing enough for further development of the concept of CRE.

Rao etal. (2004) suggested the Cumulative Residual Entropy (CRE), which is the extension of the Shannon entropy to the cumulative distribution function. Considering the complementary cumulative distribution function instead of the probability density function in the definition of Shannon entropy leads to a new entropy measure named cumulative residual entropy. CRE is defined as

$$CRE(x) = -\int_{0}^{\infty} \overline{F}(x)\log[\overline{F}(x)]dx$$

Clearly, this definition is valid both for a discrete or an absolutely continuous random variable. In addition, unlike Shannon entropy it is always positive, while preserving many interesting properties of Shannon entropy. The concept of CRE has found nice interpretations and applications in the fields of reliability.

4 GAMMA FRAILTY MODEL:-

There are many applications of the gamma frailty model. We have studied the expulsion of intrauterine contraceptive devices. Also we have studied recidivism among criminals using gamma-Weibull model. Used the gamma frailty model to check the proportional hazards assumptions in his study of malignant melanoma. A formal of the goodness-of-fit tests for the gamma frailties was constructed . They also construct a new class of frailty models that extend the gamma frailty model by using certain polynomial expansions that are orthogonal with respect to the gamma density. For that extended family, they obtained an explicit expression for the marginal likelihood of the data. The order selection test is based on finding the best fitting model in such a series of expanded models. A bootstrap was used to obtain p-values for the tests. Simulations and data examples illustrated the tests performance. considered gamma distribution as frailty distribution and the log- logistic distribution as baseline distribution for bivariate survival times. Because this distribution has the advantage of having simple algebraic expressions for its survivor and hazard functions and a closed form for its distribution function. Then, they used exponential, Weibull and log-logistic as baseline hazard functions and the gamma as well as inverse Gaussian for the frailty distributions and then based on AIC criteria, all models were compared for their performance.

4.1 Gamma Frailty Model

Let a continuous random variable Z follows a gamma distribution. For identifiability, we assume Z has expected value equal to one which introduces restriction on the scale and the shape parameters. Under the restriction, the density function and the Laplace transformation of a gamma distribution reduces to,

$$f_z(z) = \begin{cases} \frac{\left(\frac{1}{\theta}\right)^{\frac{1}{\theta}}}{\Gamma\left(\frac{1}{\theta}\right)} z^{\frac{1}{\theta}-1} e^{-\frac{x}{\theta}} & ; \ z > 0, \theta > 0 \\ 0 & \end{cases} \quad (4.1)$$

and $L_s(s) = (1 + \theta s)^{-\frac{1}{\theta}}$ with variance of Z as θ. Hence by using 1.3 changes to different cumulative hazard function, we can get different survival function. Hence we can obtain different probability density function.

The survival function of gamma frailty model is given by

$$S(t) = [1 + \theta H(t)]^{-\frac{1}{\theta}}$$

Whare H(t) is cumulative hazard function.

4.2 GAMMA FRAILTY EXPONENTIAL DISTRIBUTION:-

The probability density function of exponential distribution is given by

$$f(x) = \lambda e^{-\lambda x} \quad ; \ x \geq 0, \ \lambda > 0$$

Now cumulative hazard function of exponential distribution is given by

H(t)= λt

Now we have a new survival function which can be written as

S(t) = $[1 + \theta \mathrm{H(t)}]^{-\frac{1}{\theta}}$

Then the gamma frailty exponential survival function is

S(t) = $[1 + \theta\lambda t]^{-\frac{1}{\theta}}$

Now the probability density function of gamma frailty exponential distribution is

f(t) = $-\frac{d(s(t))}{dt}$

f(t) = $\lambda\,[1\,+\,\theta\lambda t\,]^{-\frac{1}{\theta}-1}$ $;t \geq 0\,,\theta > 0, \lambda > 0$

(1) Shannon's entropy of gamma frailty exponential distribution:-

Shannon (1948) was the first to introduce entropy, known as Shannons entropy or Shannons information measure, into information theory. Shannon defined a formal measure of entropy, called Shannon entropy. For an absolutely continuous random variable X having probability density function f, Shannons entropy is defined as

$$H(t) = -\int_0^{\infty} f(t)\ \log(f(t))dt$$

Or

$$H(t) = -\int_0^{\infty} \log[\,\lambda\,(1\,+\,\theta\lambda t\,)^{-\frac{1}{\theta}-1}]\ \lambda\,[1\,+\,\theta\lambda t\,]^{-\frac{1}{\theta}-1}dt$$

Or

H(t)=-logλ+(1+θ)

(2) Renyi entropy of gamma frailty exponential distribution:-

The parametric family of entropies was introduced by Alfred Renyi in the mid 1950s as a mathematical generalization of [17]. Renyi wanted to find the most general class of information measure that preserved the additivity of statistically independent systems and were compatible with Kolmogorovs probability axioms. Renyis entropy can be defined for continuous random variables. Let f(x) be the continuous PDF then the integrated probability is

$$H(p) = \frac{1}{1-p}\log\int_{-\infty}^{\infty}[\,f(x)]^p dx$$

This is very similar to the Shannon case, showing that the differential Renyis entropy can be negative for p1. Indeed log(n) can be thought as the entropy of the uniform distribution, and so the continuous entropy is the gain obtained by substituting the uniform distribution

by the experimental samples .

$$H_p(t) = \frac{1}{1-p} \log \int_0^\infty [\lambda\,(1\ +\ \theta\lambda t\,)^{-\frac{1}{\theta}-1}]^p dt$$

Or

$H_p(t) = \frac{1}{1-p}\log[\frac{\lambda^{p-1}}{\theta - p(1+\theta)}]$; For p ≥ 1

Quadratic Renyi's entropy:-

Quadratic renyi's entropy for continuous random reads, p = 2 for quadratic Renyi entropy

$H_2(t)=-\log[\int_0^\infty (f(t))^2 dt]$

or

$H_2(t)=-\log[\frac{\lambda}{\theta+2}]$

(3) Past entropy of gamma frailty exponential distribution:-

Past entropy is given by

$H(t)= -\int_t^\infty \frac{f(x)}{F(t)} \log\left(\frac{f(x)}{F(t)}\right) dx$

$H(t)= -\int_t^\infty \frac{\lambda(1+\theta\lambda x)^{-\frac{1}{\theta}-1})}{(1+\theta\lambda t)^{-\frac{1}{\theta}}} \log\left(\frac{\lambda(1+\theta\lambda x)^{-\frac{1}{\theta}-1})}{(1+\theta\lambda t)^{-\frac{1}{\theta}}}\right) dx$

$H(t)= -\frac{1}{(1+\theta\lambda t)^{-\frac{1}{\theta}}} \int_t^\infty \lambda(1+\theta\lambda x)^{-\frac{1}{\theta}-1})\ \log\left(\lambda(1+\theta\lambda x)^{-\frac{1}{\theta}-1})\right) dx +$

$$\frac{log(1+\theta\lambda t)^{-\frac{1}{\theta}}}{(1+\theta\lambda t)^{-\frac{1}{\theta}}} \int_t^\infty \lambda(1+\theta\lambda x)^{-\frac{1}{\theta}-1})\, dx$$

$\Rightarrow -\frac{1}{1+\theta\lambda t^{-\frac{1}{\theta}}} \int_t^{\infty} \lambda(1+\theta\lambda x)^{-\frac{1}{\theta}-1} \log\left(\lambda(1+\theta\lambda x)^{-\frac{1}{\theta}-1}\right) dx$ =

$(1+\theta)[1-\log\lambda(1+\theta\lambda t)^{-\frac{1}{\theta}}] + \theta\log\lambda$

$\Rightarrow \frac{\log(1+\theta\lambda t)^{-\frac{1}{\theta}}}{(1+\theta\lambda t)^{-\frac{1}{\theta}}} \int_t^{\infty} \lambda\,(1+\theta\lambda t)^{-\frac{1}{\theta}-1}\, dx = \log(1+\theta\lambda t)^{-\frac{1}{\theta}}$

Hence,

H(t)=$(1+\theta)+\log(\frac{1+\theta\lambda t}{\lambda})$

(4) Cumulative Residual Entropy of gamma frailty exponential distribution:-

Considering the complementary cumulative distribution function instead of the probability density function in the definition of Shannon entropy leads to a new entropy measure named cumulative residual entropy (CRE) ([18]) and it is given by

CRE(x)= $-\int_{-\infty}^{\infty} \bar{F}\,(x)\log(\bar{F}(x))\;dx$

CRE(x)= $-\int_{-\infty}^{\infty}(1+\theta\lambda x)^{-\frac{1}{\theta}}\log((1+\theta\lambda t)^{-\frac{1}{\theta}})dx$

Or

CRE(x)= $\frac{1}{\lambda(1-\theta)^2}$

4.3 GAMMA FRAILTY WEIBULL DISTRIBUTION

The Weibull distribution is one of the most important, desirable and widely used lifetime distributions. It has been used in many different fields with many applications. The CDF of the Weibull distribution is simple and has a closed form which gives a simple expression of its survival and hazard functions. It is a flexible distribution that can be used to fit different kinds of of lifetime data sets in different fields. Moreover, it has a physical meaning and interpretations of its parameters. The two-parameter Weibull distribution is specified by the cumulative distribution function CDF

$F(t) = 1 - e^{-\alpha t^{\lambda}}$; x>0

Where $\alpha > 0$ and $\lambda > 0$ are the scale and shape parameters, respectively.

The corresponding hazard function HF is

$$H(x) = \alpha\lambda t^{\lambda-1} \qquad ; x > 0$$

which can be increasing, decreasing or constant depending on λ> 1, λ < 1 or λ= 1. Unfortunately, it dose not exhibit any kind of non-monotonic hazard rate shape.

Cumulative hazard function of weibull distribution is given by

$$H(t) = \alpha t^{\lambda} \qquad ; x > 0$$

Now the gamma frailty weibull survival function is given by

$$S(t) = [1 + \theta H(t)]^{-\frac{1}{\theta}}$$

or

$$S(t) = [1 + \theta\alpha t^{\lambda}]^{-\frac{1}{\theta}}$$

Now the probability density function of gamma frailty weibull distribution is written as

f(t) = $-\frac{d(S(t))}{dt}$

f(t)= $\lambda\alpha[1 + \theta\alpha t^{\lambda}]^{-\frac{1}{\theta}-1} t^{\lambda-1} \qquad ; t > 0, \theta > 0, \lambda > 0, \alpha > 0$

(1) Shannon's entropy of gamma frailty weibull distribution:-

Shannon entropy is the key concept of information theory. The concept of Shanon entropy shares some intuition with Boltzmanns and some of the mathematics developed in information theory turns out to have relevance in statistical mechanics. It has found wide applications in different fields of science and technology. It is a characteristic of probability distribution providing a measure of uncertainty associated with the probability distribution. There are different approaches to the derivation of Shannon entropy based on different postulates or axioms.

$$H(t) = -\int_0^{\infty} f(t)\log(f(t))\,dt$$

or

$H(t) = -\int_0^\infty log[\lambda\alpha(1+\theta\alpha t^\lambda)^{-\frac{(1+\theta)}{\theta}} t^{\lambda-1}]\, \lambda\alpha(1\theta\alpha t^\lambda)^{-\frac{(1+\theta)}{\theta}} t^{\lambda-1} dt$

Or

$$H(t) = -\log(\lambda\alpha) + (1+\theta)$$

(3) Past entropy of gamma frailty weibull distribution:-

Frailty models account for the clustering present in event time data. A proportional hazards model with shared frailties expresses the hazard for each subject. Often a one-parameter gamma distribution is assumed for the frailties. In this paper, we construct formal goodness-of-fit tests to test for gamma frailties. We construct a new class of frailty models that extend the gamma frailty model by using certain polynomial expansions that are orthogonal with respect to the gamma density. For this extended family, we obtain an explicit expression for the marginal likelihood of the data. The order selection test is based on finding the best fitting model in such a series of expanded models. A bootstrap is used to obtain p-values for the tests. Simulations and data examples illustrate the tests performance. Past entropy is given by

$$H(t) = -\int_t^\infty \frac{f(x)}{\overline{F}(t)} \log\left(\frac{f(x)}{\overline{F}(t)}\right) dx$$

or

$$H(t) = -\int_t^\infty \frac{\lambda\alpha(1+\theta\alpha x^\lambda)^{-\frac{1}{\theta}-1} x^{\lambda-1}}{(1+\theta\alpha t^\lambda)^{-\frac{1}{\theta}}} \log[\frac{\lambda\alpha(1+\theta\alpha x^\lambda)^{-\frac{1}{\theta}-1} x^{\lambda-1}}{(1+\theta\alpha t^\lambda)^{-\frac{1}{\theta}}}]\, dx$$

or

$$H(t) = -\frac{1}{(1+\theta\alpha t^\lambda)^{-\frac{1}{\theta}}} \int_t^\infty \lambda\alpha\left(1+\theta\alpha x^\lambda\right)^{-\frac{1}{\theta}-1} x^{\lambda-1} \log\left[\lambda\alpha\left(1+\theta\alpha x^\lambda\right)^{-\frac{1}{\theta}-1} x^{\lambda-1}\right] dx + \frac{\log(1+\theta\alpha t^\lambda)^{-\frac{1}{\theta}}}{(1+\theta\alpha t^\lambda)^{-\frac{1}{\theta}}} \int_t^\infty \lambda\alpha(1+\theta\alpha x^\lambda)^{-\frac{1}{\theta}-1} x^{\lambda-1}\, dx$$

or

$$H(t) = -\frac{1}{(1+\theta\alpha t^\lambda)^{-\frac{1}{\theta}}}\left(\frac{1}{1-\theta}\left(\frac{\alpha t^\lambda}{\theta}\right)^{-\frac{1}{\theta}}\right)\left(\frac{\alpha t^\lambda}{\theta^2} hypergeometric2F1(1,1-\theta,2-\theta,\frac{\alpha t^\lambda}{\theta})\right)$$

$$+\left(\frac{1}{\theta}-1\right)(-\theta)\left(1+2Floor\left(\frac{Arg(-\alpha t^\lambda/\theta)}{2\Pi}\right)\Pi(\alpha t^\lambda/\theta)^\theta\right)+(1+\theta)+\log\left(\frac{1+\theta\alpha t^\lambda}{\alpha\lambda}\right)$$

Hypergeometric2F1 function:-

hypergeometric2F1(a; b; c; z) = $\sum_{k=0}^{\infty}\frac{(a)_k(b)_k z^k}{(c)_k!}$ /; mod z <1 mod z=1 re(c-a-b)>0 for mod z<1 and generic parameter a,b,c, the hypergeometric2F1(a; b; c; z) is defined by the above infinite sum (that is convergent). Out side of the unit circle mod z < 1 the function 2F1(a; b; c; z) is define as the analytic continuations with respect to z of this sum with the parameters a; b; c held fixed.

Floor function:

In mathematics and computer science, the floor function map a real number to the greatest preceding integer. More precisely, Floor(x) = $\lfloor x \rfloor$ is the greatest integer less than or equal to x .In the following formulas, x and y are real numbers, k, m, and n are integers, and

Z is the set of integers (positive, negative, and zero). Floor may be defined by the set Equations $\lfloor x \rfloor = \max\{m \in Z | \, m \le x\}$, Since there is exactly one integer in a half-open interval of length one, for any realx there are unique integers m and n satisfying

x-1 < m ≤ x ≤ n < x+1 then $\lfloor x \rfloor$=m may also be taken as the definition of floor.

Examples,

X=2, $\lfloor x \rfloor$=2

x = 2:9, $\lfloor x \rfloor$= 2

x = -2:7, $\lfloor x \rfloor$ = -3

5 INVERSE GAUSSION FRAILTY MODEL:-

Frailty models are extensively used in the survival analysis to account for the unobserved heterogeneity in individual risks to disease and death. To analyze the bivariate data on related survival times (e.g. matched pairs experiments, twin or family data), the shared frailty models were suggested. The frailty model is a random effect model for time to event data which is an extension of the Cox's proportional hazards model. Bivariate survival data arises whenever each study subjects experience two events. Particular examples include failure times of paired human organs, (e.g. kidneys, eyes, lungs, breasts, etc.) and the first and the second occurrences of a given disease. In the medical literature, several authors considered paired organs of an individual as a two-component system, which work under interdependency circumstances. In industrial applications, these data may come from systems whose survival depend on the survival of two similar components.

Hence by using 1.3 changes to different cumulative hazard function, we can get different survival function. Hence we can obtain different probability density function.

Now survival function is

$$S(t) = Exp\left[\frac{1-(1+2\theta H(t))^{\frac{1}{2}}}{\theta}\right]$$

5.1 INVERSE GAUSSION FRAILTY EPONENTIAL DISTRIBUTION:-

The survival function of exponential distribution is given by

$$S(t) = Exp\left[\frac{1-(1+2\theta H(t))^{\frac{1}{2}}}{\theta}\right]$$

Cumulative hazard function of inverse Gaussion exponential distribution is given by

H(t) = λ t

Now survival function is

$S(t) = \exp[\frac{1-(1+2\theta\lambda t)^{\frac{1}{2}}}{\theta}]$

Probability density function of inverse Gaussion frailty exponential distribution is

$f(t) = \frac{\lambda}{(1+2\theta\lambda t)^{\frac{1}{2}}} \exp[\frac{1-(1+2\theta\lambda t)^{\frac{1}{2}}}{\theta}]$;$t\geq0;\lambda,\theta>0$

(1) Shannon's entropy of inverse Gaussion frailty exponential distribution:-

Shannon's entropy of inverse Gaussion frailty exponential distribution is given by

$H(t) = E[-\log(f(t))]$

$H(t) = E[-\log(\frac{\lambda}{(1+2\theta\lambda t)^{\frac{1}{2}}} \exp[\frac{1-(1+2\theta\lambda t)^{\frac{1}{2}}}{\theta}])]$

Or

H(t) = -E[log($\frac{\lambda}{(1+2\theta\lambda t)^{\frac{1}{2}}}$)]-E[$\frac{1-(1+2\theta\lambda t)^{\frac{1}{2}}}{\theta}$])]

Or

-E[log($\frac{\lambda}{(1+2\theta\lambda t)^{\frac{1}{2}}}$)] = -logλ+$e^{\theta}(\text{Eulergamma} + \log\theta)$

And

-E[$\frac{1-(1+2\theta\lambda t)^{\frac{1}{2}}}{\theta}$])]=1

Hence

H(t) = 1- log λ +$e^{\theta}(\text{Eulergamma} + \log\theta)$

Where:-

Eulergamma:-

The gamma function is defined for all complex numbers except the non-positive integers. For complex numbers with a positive real part, it is defined via a convergent improper integral:

r(z)= $\int_0^{\infty} x^{z-1}e^{-x}\, dx$

(2) Renyi's entropy of inverse Gaussion frailty exponential distribution:-

Reniy's entropy is given by

$$H(p) = \frac{1}{(1-p)} log \int_0^{\infty} [f(x)]^p\, dx$$

Or

$$H(p) = \frac{1}{1-p} log \int_0^{\infty} \left[\frac{\lambda}{(1+2\theta\lambda t)^{\frac{1}{2}}} \exp\left(\frac{1-(1+2\theta\lambda t)^{\frac{1}{2}}}{p}\right)\right]^p dx$$

Or

$$H(p) = \frac{1}{1-p}\log\left[\theta\,\lambda^{p-1}e^{\frac{p}{\theta}}\frac{\Gamma(2-p)}{(\frac{p}{\theta})^{2-p}}\right]$$

Quadratic Renyi's entropy:-

Quadratic renyi's entropy for continuous random reads ,p = 2 for quadratic Renyi entropy

$$H(2) = -\log[\int_0^\infty [f(t)]^2\,dt]$$

Or

$$H(p) = \frac{1}{1-p}log\int_0^\infty\left[\frac{\lambda}{(1+2\theta\lambda t)^{\frac{1}{2}}}\exp\left(\frac{1-(1+2\theta\lambda t)^{\frac{1}{2}}}{p}\right)\right]^p\,dx$$

$$H(2) = -\log[\theta\,\lambda e^{\frac{2}{\theta}}]$$

(4) Cumulative Residual Entropy of inverse Gaussion frailty exponential distribution:-

Cumulative Residual Entropy of inverse Gaussion frailty exponential distribution is given by

$$CRE(x) = -\int_{-\infty}^{\infty}\bar{F}(x)\log(\bar{F}(x))dx$$

$$CRE(x) = -log\int_{-\infty}^{\infty}\exp\left(\frac{1-(1+2\theta\lambda t)^{\frac{1}{2}}}{p}\right)\left(\frac{1-(1+2\theta\lambda t)^{\frac{1}{2}}}{p}\right)dx$$

$$CRE(x) = \frac{1+2\theta}{\lambda}$$

5.2 INVERSE GAUSSION FRAILTY WEIBULL DISTRIBUTION:-

The survival function of new modified weibull distribution is given by

$$S(t) = \exp\left(\frac{1-(1+2\theta H(t))^{\frac{1}{2}}}{\theta}\right)$$

Cumulative hazard function of exponential distribution is given by

$$H(t) = \alpha\, t^{\lambda}$$

Now survival function is

$$S(t) = \exp\left(\frac{1-\left(1+2\theta\alpha t^{\lambda}\right)^{\frac{1}{2}}}{\theta}\right)$$

Probability density function of inverse gaussion frailty exponential distribution is

$$f(t) = \frac{\alpha\lambda t^{\lambda-1}}{(1+2\theta\alpha t^{\lambda})}\exp\left(\frac{1-\left(1+2\theta\alpha t^{\lambda}\right)^{\frac{1}{2}}}{\theta}\right) \quad ; t \geq 0;\ \lambda, \theta, \alpha > 0$$

(1)Shannon's entropy of inverse Gaussion frailty exponential distribution:-

Shannon's entropy of inverse Gaussion frailty exponential distribution is given by

$$H(t) = E\left[-\log\left(f(t)\right)\right]$$

$$H(t) = E\left[-\log\left[\frac{\alpha\lambda t^{\lambda-1}}{(1+2\theta\alpha t^{\lambda})}\exp\left(\frac{1-\left(1+2\theta\alpha t^{\lambda}\right)^{\frac{1}{2}}}{\theta}\right)\right]\right]$$

$$H(t) = -E\left[\log\left(\frac{\alpha\lambda t^{\lambda-1}}{(1+2\theta\alpha t^{\lambda})}\right)\right] - E\left[\frac{1-\left(1+2\theta\alpha t^{\lambda}\right)^{\frac{1}{2}}}{\theta}\right]$$

$$\begin{aligned} -E\left[\log\left(\frac{\alpha\lambda t^{\lambda-1}}{(1+2\theta\alpha t^{\lambda})}\right)\right] \\ &= -\log(\alpha\lambda) + \frac{1-\lambda}{\theta\lambda}\left[\frac{1}{\theta}\left(\text{Eluergamma } + \log\left(\frac{1}{\theta}\right)\right)\right] \\ &+ \frac{1-\lambda}{\theta\lambda}\left[\frac{1}{\theta}\left(e^{2\theta}\text{incompletegamma}(0,2\theta) + \log 2\right)\right] \quad + \frac{1-\lambda}{\theta}\log(2\theta\alpha) \\ &+ e^{\theta}\left[(\text{incompletegamma}(0,\theta))\right] \end{aligned}$$

and

$$-E\left[\frac{1-\left(1+2\theta\alpha t^{\lambda}\right)^{\frac{1}{2}}}{\theta}\right]=-\frac{1}{\theta}+e^{\frac{1}{\theta}}$$

$$H(t)=-\log(\alpha\lambda)+\frac{1-\lambda}{\theta\lambda}\left[\frac{1}{\theta}\left(\text{Eluergamma }+\log\left(\frac{1}{\theta}\right)\right)\right]$$
$$+\frac{1-\lambda}{\theta\lambda}\left[\frac{1}{\theta}\left(e^{2\theta}\text{incompletegamma}(0,2\theta)+\log 2\right)\right]\quad+\frac{1-\lambda}{\theta}\log(2\theta\alpha)$$
$$+e^{\theta}[(\text{incompletegamma}(0,\theta))]-\frac{1}{\theta}+e^{\frac{1}{\theta}}$$

where:-

Incompletegamma:-

The "complete" gamma function Gamma(a) can be generalized to the incomplete gamma function Gamma(a,x) such that Gamma(a)=Gamma(a,0). This "upper" incomplete gamma function is given by

Gamma(a,x)= $\int_{z}^{\infty} t^{a-1}e^{-t}dt$

The lower incomplete gamma function is given by

Gamma(a,x)= $\int_{0}^{x} t^{a-1}e^{-t}dt$

The incomplete gamma function Gamma(0,x) has continued fraction

Gamma(0,x) = $\frac{e^{-x}}{x+1-\frac{1}{x+3-\frac{4}{x+5-\frac{9}{x+7-\ldots\ldots\ldots\ldots}}}}$

REFERENCES

1. Shannon, C.E., 1948. A mathematical theory of communications. Bell System Tech. J. 27, 379-423.
2. Rnyi, A., 1961. On measures of entropy and information. Proceedings of the fourth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, pp. 547-561.
3. Arimoto, S., 1971. Information-theoretical considerations on estimation problems. In- form. Control 19, 181-194. Belzunce, F., Navarro, J., Ruiz, J.M., Aguila, Y., 2004. Some results on residual entropy function. Metrika 59, 147-161.

4. Ferreri, C., 1980. Hypoentropy and related heterogeneity divergence measures. Statis-tica 40, 55-118.

5. Havrda, J., Charvt, F., 1967. Quantification method of classification process: concept of structural entropy. Kybernetika 3, 30-35.

6. Khinchin, A.J., 1957. Mathematical Foundation of Information Theory. Dover, NewYork.

7. Sharma, B.D., Mittal, P., 1977. New non-additive measures of relative information. J. Combin, Inform. System

8. Sci. 2, 122133. Sharma, B.D., Taneja, I.J., 1975. Entropy of type and other generalized measures in information theory. Metrika 22, 205-215.

9. Ebrahimi, N., (1996) How to measure uncertainty in the residual life distributions. Sankhya Ser.A 58, 48-56.

10. Ebrahimi, N., Kirmani, S.N.U.A.,(1996) Some results on ordering of survival functions through uncertainty. Statist.Prob.Lett.11,167-176.

11. Di Crescenzo, A., Longobardi, M., 2002. Entropy-based measure of uncertainty in past lifetime distributions. J. Appl. Probab. 39, 434440.

12. J.W. Vaupel, K.G. Manton, and E. Stallard, The impact of heterogeneity in individual frailty on the dynamics of mortality, Demography 16, 1979, 439-454.

13. T. Lancaster, Analysis of Transition Data (University of Cambridge, New work, 1990).

14. P. Hougaard, Life Table Methods for Heterogeneous Populations: Distributions De- scribing Heterogeneity.Biometrika,71, 1984,75-83.

15. Belzunce F., Guillamon, A., Navarro, J., Ruiz, J.M. (2001) Kernel estimation of residual entropty. Commun. Statist. - Theory Meth., 30(7), 1243-1255.

16. C. E. Shannon, A mathematical theory of communication, The Bell System Technical Journal, vol. 27, no. 3, pp. 379423, 1948. 25

17. C. Adami. Information theory in molecular biology. Physics of Life Reviews, 1:322,