Generalized Measures of Divergence for Lifetime Distributions

Ilia Vonta
Department of Mathematics
National Technical University of Athens
Athens, GREECE
vonta@math.ntua.gr

Abstract

Measures of divergence or discrepancy are used either to measure mutual information concerning two variables or to construct model selection criteria. In this paper we are focusing on divergence measures that are based on a class of measures known as Csiszar's divergence measures. In particular, we propose a measure of divergence between residual lives of two items that have both survived up to some time t as well as a measure of divergence between past lives, both based on Csiszar's class of measures. Furthermore, we derive properties of these measures and provide examples based on the Cox model and frailty or transformation models.

1 Introduction

A measure of divergence is used as a way to evaluate the distance (divergence) between any two populations or functions. In the present work, we concentrate on divergence measures that are based on a class of measures known as Csiszar's family of divergence measures or Csiszar's φ-divergence (Csiszar, (1963); Ali and Silvey, (1966)).

An issue of fundamental importance in Statistics is the investigation of Information Measures. These measures are classified in different categories and measure the quantity of information contained in the data with respect to a parameter θ, the divergence between two populations or functions, the information we get after the execution of an experiment and other important information according to the application they are used for. Traditionally, the measures of information are classified in four main categories namely divergence - type, entropy - type, Fisher - type and Bayesian - type.

Measures of divergence between two probability distributions have a very long history initiated by the pioneer work of Pearson, Mahalanobis, Lévy and Kolmogorov. Among the most popular measures of divergence are the Kullback-Leibler measure of divergence and the Csiszar's φ-divergence family of measures. Recently, the BHHJ divergence measure was proposed by Basu et *al.* (1998) and generalized to the BHHJ family of measures by Mattheou et *al.* (2009).

Ebrahimi and Kirmani (1996a) introduced a measure of discrepancy between the lifetimes X and Y of two items at time t. In survival analysis or in reliability we might know the current age t of a biomedical or technical system. We need to take this information into consideration when we compare two systems or populations. Ebrahimi and Kirmani (1996a) achieved this by replacing the distribution functions of the random variables X and Y in the Kullback-Leibler divergence of X and Y, by the distributions of their residual lifetimes. Di Crescenzo and Longobardi (2004) define a dual measure of divergence which constitutes a distance between past life distributions.

2 Generalized Measures of Divergence

Let X and Y be absolutely continuous, non-negative random variables that describe the lifetimes of two items. Let $f(x)$, $F(x)$ and $\overline{F}(x)$ be the density function, the cumulative distribution function and the survival function of X respectively. Let also $g(x)$, $G(x)$ and $\overline{G}(x)$ be the density function, the cumulative distribution function and the survival function of Y respectively. Let $h_X(x) = f(x)/\overline{F}(x)$ and $h_Y(x) = g(x)/\overline{G}(x)$ be the hazard rate functions of X and Y while $\tau_X(x) = f(x)/F(x)$ and $\tau_Y(x) = g(x)/G(x)$ are the reversed hazard rate functions of X and Y. Without loss of generality we assume throughout the paper that the support of f and g is $(0, +\infty)$.

The Kullback-Leibler distance between F and G is defined by

$$I_{X,Y} = \int_0^\infty f(x) \log\left(\frac{f(x)}{g(x)}\right) dx \tag{1}$$

where log denotes the natural logarithm. A generalization of this distance is defined as

$$I_{X,Y}^{\varphi} = \int_0^\infty g(x) \varphi\left(\frac{f(x)}{g(x)}\right) dx \tag{2}$$

and is known as Csiszar's family of measures of divergence.

When the function φ is defined as $\varphi(u) = u \log u$ or $\varphi(u) = u \log u + 1 - u$ then the above measure reduces to the Kullback-Leibler measure. If $\varphi(u) = (1-u)^2$, Csiszar's measure yields the Pearson's chi-square divergence. If $\varphi(u) = \left(u^{a+1} - u - a(u-1)\right)/(a(a+1))$ we obtain the Cressie and Read power divergence (Cressie and Read, (1984)), $a \neq 0, -1$. If $\varphi(u) = (1-\sqrt{u})^2$, we obtain the Matusita's divergence (Matusita, (1967)).

We define also the function

$$\varphi(u) = 1 - (1 + \frac{1}{a})u + \frac{u^{1+a}}{a}, \quad a > 0$$

which is related to a recently proposed measure of divergence (BHHJ power divergence, Basu et *al.* (1998)). Another function that we consider is

$$\varphi(u) = u^{1+a} - (1 + \frac{1}{a})u^a + \frac{1}{a}, \quad a > 0.$$

These last two functions are special cases of the BHHJ family of measures of divergence proposed by Mattheou et *al.* (2009)

$$I_X^a(g, f) = E_g\left(g^a(X)\varphi\left(\frac{f(X)}{g(X)}\right)\right) = \int g^{1+a}(z)\, \varphi\left(\frac{f(z)}{g(z)}\right) d\mu, \quad a > 0, \tag{3}$$

where μ represents the Lebesgue measure. Appropriately chosen functions $\varphi(\cdot)$ give rise to special measures mentioned above.

Ebrahimi and Kirmani (1996a) introduced a measure of discrepancy between X and Y at time t as follows

$$I_{X,Y}(t) = \int_t^\infty \frac{f(x)}{\overline{F}(t)} \log\left(\frac{f(x)/\overline{F}(t)}{g(x)/\overline{G}(t)}\right) dx, \qquad t > 0. \tag{4}$$

A dual measure is defined in Di Crescenzo and Longobardi (2004) which constitutes a distance between past lifetimes

$$\overline{I}_{X,Y}(t) = \int_0^t \frac{f(x)}{F(t)} \log\left(\frac{f(x)/F(t)}{g(x)/G(t)}\right) dx, \qquad t > 0. \tag{5}$$

In this paper we propose two measures of discrepancy which are based on the Csiszar's φ-divergence family, namely, the *φ-distance between residual lifetimes*

$$I_{X,Y}^{\varphi}(t) = \int_t^\infty \frac{g(x)}{\overline{G}(t)}\, \varphi\left(\frac{f(x)/\overline{F}(t)}{g(x)/\overline{G}(t)}\right) dx, \qquad t > 0 \tag{6}$$

and the *φ-distance between past lifetimes*

$$\overline{I}_{X,Y}^{\varphi}(t) = \int_0^t \frac{g(x)}{G(t)} \varphi\left(\frac{f(x)/F(t)}{g(x)/G(t)}\right) dx, \qquad t > 0. \tag{7}$$

where the function φ belongs to a class of convex functions Φ that satisfy some regularity conditions.

3 Measures in Survival Analysis and Reliability Models

3.1 Proportional hazards and proportional reverse hazards models

In this section we examine properties of the proposed measures of divergence and find various discrimination measures in cases like the proportional hazards model, the proportional reverse hazards model and the frailty or transformation models. For the latter case, we provide the φ-distance between the respective residual and past lifetimes associated with the Cox and frailty models respectively.

For the case of proportional hazards let X and Y be random variables with distribution functions F and G respectively for which it holds that

$$\overline{G}(x) = (\overline{F}(x))^{\theta} \text{ for all } x > 0 \text{ and } \theta > 0. \tag{8}$$

Theorem 1. *The discrimination measure $I^{\varphi}_{X,Y}(t)$ between the random variables X and Y which satisfy the proportional hazards assumption (8) is independent of t and is given as*

$$I^{\varphi}_{X,Y}(t) = \int_0^1 \varphi\left(\frac{1}{\theta y^{\theta-1}}\right) dy^{\theta} \tag{9}$$

(ii) If $I^{\varphi}_{X,Y}(t)$ is independent of t, then there exists a constant $\theta > 0$ such that (8) holds.

Let now X and Y be random variables with distribution functions F and G respectively which satisfy the proportional hazards assumption but with reverse proportionality, that is, $\overline{F}(x) = (\overline{G}(x))^{\theta}$. In this case, the discrimination measure takes the form

$$I^{\varphi}_{X,Y}(t) = \int_0^1 \varphi\left(\theta y^{\theta-1}\right) dy. \tag{10}$$

For the proportional reverse hazards model which is defined as

$$G(x) = (F(x))^{\theta} \text{ for all } x > 0 \text{ and } \theta > 0 \tag{11}$$

the following result holds.

Theorem 2. *(i) The discrimination measure $\overline{I}^{\varphi}_{X,Y}(t)$ between the random variables X and Y which satisfy the proportional reverse hazards assumption (11) is independent of t and is given as*

$$\overline{I}^{\varphi}_{X,Y}(t) = \int_0^1 \varphi\left(\frac{1}{\theta y^{\theta-1}}\right) dy^{\theta} \tag{12}$$

(ii) If $\overline{I}^{\varphi}_{X,Y}(t)$ is independent of t, then there exists a constant $\theta > 0$ such that (11) holds.

Let now X and Y be random variables with distribution functions F and G respectively which satisfy the proportional reverse hazards assumption but with reverse proportionality. In that case, the discrimination measure $\overline{I}^{\varphi}_{X,Y}(t)$ is given by (10).

3.2 Frailty or transformation model versus Cox proportional hazards model

Let now X and Y be random variables with distribution functions F_1 and F_2, probability density functions f_1 and f_2 and survival functions S_1 and S_2 respectively. Let H be the baseline cumulative hazard function and h the baseline intensity hazard function. Let X follow a Cox model (Cox 1972) under which

$$S_1(x) = e^{-\theta H(x)}, \quad \theta > 0. \tag{13}$$

Let also Y follow a frailty model under which (Vonta 1996)

$$S_2(x) = e^{-G(\theta H(x))}, \quad \theta > 0 \tag{14}$$

where the function G is assumed to be concave, increasing with $G(0) = 0$ and $G(\infty) = \infty$.

We have derived the usual divergence between the distributions of X and Y using the Kullback-Leibler and the Csiszar's divergence as well as the φ-distance between the respective past and residual lifetimes of X and Y. We provide here for the latter cases the relative results which are stated below.

Theorem 3. *The discrimination measure $I^{\varphi}_{X,Y}(t)$ between the random variables X and Y which follow the Cox proportional hazards model and the frailty or transformation model (14) respectively, is given as*

$$I^{\varphi}_{X,Y}(t) = \int_{\theta H(t)}^{\infty} \frac{e^{-G(y)}G'(y)}{e^{-G(\theta H(t))}} \; \varphi\left(\frac{e^{-y}/e^{-\theta H(t)}}{e^{-G(y)}G'(y)/e^{-G(\theta H(t))}}\right) dy \tag{15}$$

for $t > 0$.

Theorem 4. *The discrimination measure $\overline{I}^{\varphi}_{X,Y}(t)$ between the random variables X and Y which follow the Cox proportional hazards model and the frailty or transformation model (14) respectively, is given as*

$$\overline{I}^{\varphi}_{X,Y}(t) = \int_{0}^{\theta H(t)} \frac{e^{-G(y)}G'(y)}{1 - e^{-G(\theta H(t))}} \; \varphi\left(\frac{e^{-y}/1 - e^{-\theta H(t)}}{e^{-G(y)}G'(y)/1 - e^{-G(\theta H(t))}}\right) dy \tag{16}$$

for $t > 0$.

References

[1] Ali, S. M. and Silvey, S. D. (1966). A general class of coefficients of divergence of one distribution from another, *J. Roy. Statist. Soc. B*, *28*, 131-142.

[2] Basu, A., Harris, I. R., Hjort, N. L. and Jones, M. C. (1998). Robust and efficient estimation by minimising a density power divergence, *Biometrika*, *85*, 549—559.

[3] Cox, D. R. (1972). Regression models and life tables, *J. Roy. Statist. Soc.*, *B 34*, 187-202.

[4] Cressie, N. and Read, T. R. C. (1984). Multinomial goodness-of-fit tests, *J. R. Statist. Soc*, *5*, 440-454.

[5] Csiszar, I. (1963). Eine Informationstheoretische Ungleichung und ihre Anwendung auf den Bewis der Ergodizitat on Markhoffschen Ketten, *Publ. of the Math. Inst. of the Hungarian Academy of Sc.*, *8*, 84–108.

[6] Di Crescenzo, A. and Longobardi, M. (2002) Entropy-based measure of uncertainty in past lifetime distributions, *J. Appl. Prob.*, *39*, 434-440.

[7] Di Crescenzo, A. and Longobardi, M. (2004) A measure of discrimination between past lifetime distributions, *Statist. Probab. Lett.*, *67*, 173-182.

[8] Ebrahimi, N. and Kirmani, S.N.U.A. (1996a) A measure of discrimination between two residual life-time distributions and its applications, *Ann, Inst. Statist. Math.*, *48*, 257-265.

[9] Ebrahimi, N. and Kirmani, S.N.U.A. (1996b) A characterization of the proportional hazards model through a measure of discrimination between two residual life distributions, *Biometrika*, *83*, 233-235.

[10] Mattheou, K., Lee, S., and Karagrigoriou, A. (2009). A model selection criterion based on the BHHJ measure of divergence, *J. of Statist. Plan. and Infer.*, *139*, 128-135.

[11] Matusita, K. (1967). On the notion of affinity of several distributions and some of its applications, *Ann. Inst. Statist. Math.*, *19*, 181-192.

[12] Vonta, F. (1996) Efficient estimation in a nonproportional hazards model in survival analysis, *Scand. Jour. Statist.*, *23*, 49-62.