Ann. Inst. Statist. Math.
Vol. 53, No. 3, 528–542 (2001)

INFORMATION IN QUANTAL RESPONSE DATA AND RANDOM CENSORING

CH. TSAIRIDIS, K. ZOGRAFOS, K. FERENTINOS AND T. PAPAIOANNOU*

Department of Mathematics, University of Ioannina, 45110 Ioannina, Greece

(Received November 5, 1998; revised March 6, 2000)

Abstract. In this paper we study interesting properties of Fisher and divergence type measures of information for quantal, complete and incomplete random censoring, and not censoring at all. It is shown that, while quantal random censoring is less expensive, it is less informative than complete random censoring. It is also shown that in experiments which are mixtures of quantal and complete random censoring, the information received from these experiments is a convex combination of quantal information and the information in complete random censoring. Finally, the "total information" property is studied, in which the information received by the uncensored experiment can be expressed as the sum of the information provided by random censoring and the loss of information due to censoring. The results for Fisher's measure of information are an extension of already known results to the multiparameter case. The investigation of the previous information properties for divergence type measures is a new element, as well as the comparison of byproducts of Fisher information matrices.

Key words and phrases: Quantal random censoring, complete random censoring, Fisher information matrix, φ-divergence, total information.

1. Introduction

In life testing experiments the response variable, the lifetime, is almost never fully observed, because there is not enough time and/or money to run the testing, until all units, put on the test, fail. This is the reason why censoring is always used in life testing. In clinical trials, where the variable of primary interest is the time to survival (overall survival, time to disease progression, etc.), censoring plays a fundamental role. Here we shall be concerned with random right censoring.

Let $X_1, X_2, \dots, X_n$ be independent and identically disributed (i.i.d.) random variables with common survival function $\bar{F}(t) = P(X > t)$, where X represents the time taken for an event of interest to occur. It is assumed the X_i's are not all observed fully, but some are censored on the right. In this case, instead of X, we observe Y, which is the censoring variable with survival function $\bar{G}(t) = P(Y > t)$. X and Y will be assumed to be independent. The recorded information (Z_1, δ_1), $(Z_2, \delta_2), \dots, (Z_n, \delta_n)$ is obtained from the random pair (Z, δ), where $Z = \min(X, Y)$ and $\delta = I_{(X \le Y)}$. Here, δ_i indicates whether X_i has been observed or not. Note that if $F = G$, then the probability of censoring equals $1/2$, and if Y is stochastically larger than X, then the probability of censoring is less than $1/2$.

*Now at Department of Statistics and Insurance Science, University of Piraeus, 18534 Piraeus, Greece.

In some situations, continuous follow up or monitoring of the tested items is very costly or technically difficult, and a failure or the item's true condition is found only on inspection. For example, a cracked part inside a machine. The machine or module must be taken apart, examined and assembled again. In other situations, failures are not signalled when they occur. To reduce the testing cost and make it simpler, we inspect the item's condition at a randomly selected inspection time Y. If a unit is found failed, one knows only that its failure time was before its inspection time. Similarly, if a unit is found unfailed, one knows that its failure time is beyond its inspection time. The information received is in the binary form "$x \leq Y$" or "$x > Y$", where x is the failure time of the item. If the lifetime $x > Y$, the information recorded is only that $x > Y$. This is the model termed by Nelson (1982) as quantal response model. Data associated with this model are generally called *quantal response data*, and the corresponding censoring mode will be termed *quantal type censoring*. If the inspection time is random, then we have *quantal random censoring*. We shall distinguish two types of quantal random censoring: (i) *quantal random censoring based on* δ, where we only observe the δ_i, $i = 1, 2, \ldots, n$, and (ii) *quantal random censoring based on* (δ, Y), where we not only observe the δ_i, but for each δ_i we also have a Y_i, the inspection time, $i = 1, 2, \ldots, n$. Both types of quantal random censoring are less expensive than complete random censoring, but less informative, as we intuitively feel. We establish this, by means of information theory in Section 2 of this paper. In other words, if I is any measure of information, and I_{rc}, I_{qrc} and I_δ are its values for random censoring and quantal random censoring based on (δ, Y) and δ, respectively, we shall prove that $I_\delta \leq I_{qrc} \leq I_{rc}$, for all standard I.

There are situations where, due to monitoring reasons, complete random censoring is not possible. For instance, an uncensored observation is either recorded with constant probability or varying probability over the experimental units; of course, the examination or the inspection time is always recorded. In other words, uncensored observations are randomly recorded or signalled. This experiment can be thought as a *mixture of quantal and (complete) random censoring*, and it turns out that information for this experiment is a convex combination of quantal information and the information in complete random censoring. This is presented in Section 3, where we prove that, if p is the probability of recording an uncensored observation, and I the information in this experiment, then $I = pI_{rc} + (1-p)I_{qrc}$, for any of the standard measures of information.

It is intuitively obvious that the information provided in a censored experiment should be less than or equal to the information provided by the uncensored experiment. In other words, if I is a measure of information, we should have $I \geq I_{rc}$. Indeed, this has been shown to be true for any measure of information I (Tsairidis *et al.* (1996), Theorem 5.3). This result leads to the idea of decomposing I into two parts, one being I_{rc}, and the other a meaningful conditional information component, which is the loss of information due to censoring. This property has been called by Gertsbakh and Kagan (1999) as the *total information* property, and it is shown to be valid for Fisher's information matrix and Kullback-Leibler measure of information. In addition, based on this property, it is shown that, if both measures of information are *equivariant* with respect to the expected value of the distribution function of lifetime (with respect to G), then the conditional information equals the unconditional, which is a kind of *lack-of-memory property*. This is the topic of discussion of Section 4 of this paper.

There are two types of measures of information: parametric or Fisher type, and non-parametric or divergence type (cf. Ferentinos and Papaioannou (1981), and Papaioannou (1985)). The main representative of the first type is Fisher's measure of information, which is defined for parametric families of distributions, while for the second class is Csiszar's φ-divergence (cf. Csiszar (1963)) or the Kullback-Leibler measure of information (cf. Kullback and Leibler (1951)). Another well known divergence is Matusita's measure of information (cf. Matusita (1967)). In random censoring we shall follow the definition of measures of information as they have been clarified in Tsairidis *et al.* (1996), and are based on the distribution of the random variables Z and δ or the likelihood for a single observation. Where needed, we shall sometimes employ a conditional expectation argument to obtain simplified expressions for the measures of information. Frequently, these expressions are analogous to the Hollander, Proschan, and Sconing type measure of information (cf. Hollander *et al.* (1987)).

The distributional assumptions are as usual. For Fisher type measures of information, we assume a parametric family of distributions $\mathcal{M} = \{P_\theta, \theta \in \Theta\}$ with g.p.d.f. $f(x;\theta) = \frac{dP_\theta}{d\mu}$, relative to a dominating measure μ on the real space $\mathcal{X}$, satisfying some appropriate regularity conditions (cf. Papaioannou (1985)). $\mathcal{X}$ will be assumed nonnegative, and the parameter space is an open subset of the Euclidean space R^k. For divergence type measures of information, we consider two probability measures P_1 and P_2 on the measurable space $(\mathcal{X}, \mathcal{T})$. Let $f_i(x) = \frac{dP_i}{d\mu}$, $i = 1, 2$, where μ is a dominating finite or σ-finite measure, and $F_i(x)$, $i = 1, 2$, be the g.p.d.f's and c.d.f's of X, respectively. For the censoring variable Y we shall assume a g.p.d.f. $g(y)$ and a c.d.f. $G(y)$, independent of θ (noninformative censoring). The distribution of Y need not belong to $\mathcal{M}$. For the φ-divergence or Csiszar's measure of information we shall assume that φ is a convex function satisfying appropriate regularity conditions (cf. Csiszar (1963)).

Measures of information can be partially ordered as follows: If W is a function of X, then $I_X \geq I_W$, for all standard I, where the ordering relationship "$\geq$" is in the positive definite sense if I is an information matrix (cf. Pukelsheim (1993)). This is the *monotonicity* property of measures of information or the well known property of *maximal information* (cf. Ferentinos and Papaioannou (1981), and Papaioannou (1985)), applied to any paired random variables X and W defined as before. This ordering carries over to experiments, and, in particular, to censored experiments. An experiment $\mathcal{E}$ is identified with a random variable X and will be denoted by $\mathcal{E}_X$. Thus, experiment $\mathcal{E}_X$ is "greater than or equal to" (or sufficient in Blackwell's (1951) sense) experiment $\mathcal{E}_W$, $\mathcal{E}_X \succeq \mathcal{E}_W$, if W is a function of X. The monotonicity property of measures of information with respect to experiments states that if $\mathcal{E}_{X_1} \succeq \mathcal{E}_{X_2}$, then $I_{X_1} \geq I_{X_2}$. The specification for censored experiments is obvious.

In summary, in this paper we study interesting properties of Fisher and divergence type measures of information for quantal random censoring, complete random censoring, incomplete random censoring, and not censoring at all. The motivation for this work have been the papers by Elperin and Gertsbakh (1988), Gertsbakh (1995), and Gertsbakh and Kagan (1999), where similar properties have been examined only for Fisher's information measure with θ univariate. Here we extend these results to Fisher's information matrix and its byproducts, and to divergence type measures of information, as well. These results might be useful in deciding which type of censoring should be preferred for designing life testing experiments or clinical trials.

2. Quantal and randomly censored information

For the random censoring setup, defined in Section 1, the joint distribution of (Z, δ) or the likelihood for a single observation, is given by

$$p((z,\delta);\theta) = [f(z;\theta)\bar{G}(z)]^{\delta}[g(z)\bar{F}(z;\theta)]^{1-\delta}, \qquad \delta = 0,1.$$

Fisher's measure of information is given by (cf. Tsairidis *et al.* (1996))

$$I_{rc}^{F}(\theta) \equiv I_{(Z,\delta)}^{F(rc)}(\theta) = \left\| \int_0^{\infty} \frac{f_{(i)}(z;\theta)f_{(j)}(z;\theta)}{f(z;\theta)}\bar{G}dz + \int_0^{\infty} \frac{\bar{F}_{(i)}(z;\theta)\bar{F}_{(j)}(z;\theta)}{\bar{F}(z;\theta)} g(z)dz \right\|_{k\times k}$$

($\|\cdot\|_{k\times k}$ denotes a $k \times k$ matrix), where

$$f_{(i)}(z;\theta) = \frac{\partial f(z;\theta)}{\partial \theta_i}, \qquad F_{(i)}(z;\theta) = \frac{\partial F(z;\theta)}{\partial \theta_i}, \quad i = 1,2,\ldots,k.$$

The following conditional expectation argument gives a simpler expression for $I_{rc}^{F}(\theta)$: The conditional distribution of (Z,δ) given $Y = y$, is given by

$$p((z,\delta) \mid y;\theta) = [f(z;\theta)]^{\delta}[\bar{F}(y;\theta)]^{1-\delta}, \qquad \delta = 0,1.$$

Note that for $\delta = 1$, $z \leq y$, and for $\delta = 0$, $z > y$. The conditional Fisher information matrix based on the above distribution, after some algebra, is

$$I_{(Z,\delta)|Y=y}^{F(rc)}(\theta) = \left\| \int_0^{y} \frac{f_{(i)}(z;\theta)f_{(j)}(z;\theta)}{f(z;\theta)}dz + \frac{F_{(i)}(y;\theta)F_{(j)}(y;\theta)}{\bar{F}(y;\theta)} \right\|_{k\times k}.$$

Unconditionally we have

$$\text{(2.1)} \quad I_{rc}^{F}(\theta) \equiv I_{(Z,\delta)}^{F(rc)}(\theta) = \left\| \int_0^{\infty} g(y) \left[\int_0^{y} \frac{f_{(i)}(z;\theta)f_{(j)}(z;\theta)}{f(z;\theta)}dz + \frac{F_{(i)}(y;\theta)F_{(j)}(y;\theta)}{\bar{F}(y;\theta)} \right] dy \right\|_{k\times k}.$$

This will be called *randomly censored Fisher information.*

Csiszar's φ-divergence between f_1 and f_2 based on the censored experiment at hand, is given by (cf. Tsairidis *et al.* (1996))

$$\begin{aligned} I_{(Z,\delta)}^{C(rc)}(f_1,f_2) &= \sum_{\delta=0,1} \int_0^{\infty} p_2(z,\delta)\varphi\left(\frac{p_1(z,\delta)}{p_2(z,\delta)}\right) dz \\ &= \int_0^{\infty} f_2(z)\bar{G}(z)\varphi\left(\frac{f_1(z)}{f_2(z)}\right) dz + \int_0^{\infty} g(z)\bar{F}_2(z)\varphi\left(\frac{\bar{F}_1(z)}{\bar{F}_2(z)}\right) dz, \end{aligned}$$

where $p_i(z,\delta)$, $i = 1,2$, is $p((z,\delta);\theta)$ with f replaced by f_i, $i = 1,2$, and $\varphi(\cdot)$ is the convex function associated with Csiszar's measure of information, satisfying appropriate regularity conditions for the existence of the above integral (cf. Csiszar (1963)). Again, the conditional expectation argument gives

$$I_{(Z,\delta)|Y=y}^{C(rc)}(f_1,f_2) = \int_0^{y} f_2(z)\varphi\left(\frac{f_1(z)}{f_2(z)}\right) dz + \bar{F}_2(y)\varphi\left(\frac{\bar{F}_1(y)}{\bar{F}_2(y)}\right),$$

and unconditionally

$$\text{(2.2)} \quad I_{rc}^{C}(f_1,f_2) \equiv I_{(Z,\delta)}^{C(rc)}(f_1,f_2) = \int_0^{\infty} g(y)\left[\int_0^{y} f_2(z)\varphi\left(\frac{f_1(z)}{f_2(z)}\right) dz + \bar{F}_2(y)\varphi\left(\frac{\bar{F}_1(y)}{\bar{F}_2(y)}\right)\right] dy.$$

This will be called *randomly censored divergence*.

In quantal random censoring we either observe only the variables δ_i or the variables (δ_i, Y_i), $i = 1, 2, \ldots, n$, in contrast to the pairs (Z_i, δ_i), $i = 1, 2, \ldots, n$, which are observed when we have plain random censoring. The distribution of δ_i is $Bernoulli(1, p = \int_0^\infty f(x;\theta)\bar{G}(x)dx)$.

In *quantal random censoring based on* δ, Fisher's information matrix will be denoted by $I_\delta^F(\theta)$, and is given by

$$I_\delta^F(\theta) = \left\| E\left(\frac{\partial \log h(\delta;\theta)}{\partial \theta_i}\frac{\partial \log h(\delta;\theta)}{\partial \theta_j}\right)\right\|_{k\times k}, \tag{2.3}$$

where

$$h(\delta;\theta) = \begin{cases} \int_0^\infty f(x;\theta)\bar{G}(x)dx \ , \ \delta = 1 \\ \\ \int_0^\infty g(x)\bar{F}(x;\theta)dx \ , \ \delta = 0 \end{cases},$$

while Csiszar's measure of information will be denoted by $I_\delta^C(f_1, f_2)$, and is given by

$$I_\delta^C(f_1, f_2) = \sum_{\delta=0,1} h_2(\delta)\varphi\left(\frac{h_1(\delta)}{h_2(\delta)}\right), \tag{2.4}$$

where $h_i(\delta)$, $i = 1, 2$, is $h(\delta)$ with f replaced by f_i, $i = 1, 2$.

For *quantal random censoring based on* (δ, Y), we shall consider the distribution of δ given $Y = y$, which is a Bernoulli distribution with probability of success $p = F(y;\theta)$, that is

$$q(\delta \mid y;\theta) = [F(y;\theta)]^\delta \left[\bar{F}(y;\theta)\right]^{1-\delta}, \quad \delta = 0, 1. \tag{2.5}$$

The conditional Fisher information matrix based on the above distribution, after some algebra, is

$$I_{(\delta,Y)|Y=y}^{F(qrc)}(\theta) = \left\|\frac{F_{(i)}(y;\theta)F_{(j)}(y;\theta)}{F(y;\theta)\bar{F}(y;\theta)}\right\|_{k\times k}.$$

Unconditionally we have

$$I_{qrc}^F(\theta) \equiv I_{(\delta,Y)}^{F(qrc)}(\theta) = \left\|\int_0^\infty g(y)\frac{F_{(i)}(y;\theta)F_{(j)}(y;\theta)}{F(y;\theta)\bar{F}(y;\theta)}dy\right\|_{k\times k}. \tag{2.6}$$

This will be called *quantal randomly censored Fisher information*.

In a similar manner, if f_i and F_i, $i = 1, 2$, are the p.d.f's and c.d.f's of X, respectively, and g and G the p.d.f. and c.d.f. of Y, respectively, then in view of (2.5), Csiszar's φ-divergence between f_1 and f_2 based on the conditional random variable $(\delta, Y) \mid Y = y$ is

$$I_{(\delta,Y)|Y=y}^{C(qrc)}(f_1, f_2) = F_2(y)\varphi\left(\frac{F_1(y)}{F_2(y)}\right) + \bar{F}_2(y)\varphi\left(\frac{\bar{F}_1(y)}{\bar{F}_2(y)}\right),$$

and unconditionally

$$\begin{aligned} I_{qrc}^C(f_1, f_2) &\equiv I_{(\delta,Y)}^{C(qrc)}(f_1, f_2) \\ &= \int_0^\infty g(y)\left[F_2(y)\varphi\left(\frac{F_1(y)}{F_2(y)}\right) + \bar{F}_2(y)\varphi\left(\frac{\bar{F}_1(y)}{\bar{F}_2(y)}\right)\right]dy. \end{aligned} \tag{2.7}$$

This will be called *quantal randomly censored divergence*.

The following theorem shows that the monotonicity property holds in the censoring case. In particular, it is shows that, both randomly censored Fisher information matrix and randomly censored divergence, are greater than or equal to their quantal censored counterparts, respectively, provided that the two modes of censoring (random censoring and quantal random censoring) have the same censoring variable Y (equal censoring times). Elperin and Gertsbakh (1988) established the validity of (i), with θ univariate.

THEOREM 2.1. *For the randomly censored and the quantal randomly censored Fisher information matrix, defined by* (2.1), (2.3) *and* (2.6), *respectively, and the randomly censored and quantal randomly censored divergence, defined by* (2.2), (2.4) *and* (2.7), *respectively, we have*

(i) $I_\delta^F(\theta) \leq I_{qrc}^F(\theta) \leq I_{rc}^F(\theta)$, *in the positive semidefinite sense, and*

(ii) $I_\delta^C(f_1, f_2) \leq I_{qrc}^C(f_1, f_2) \leq I_{rc}^C(f_1, f_2)$.

PROOF. The inequality $I_\delta \leq I_{qrc}$ follows immediately from the monotonicity or maximal information property, which is satisfied for Fisher's information matrix and φ-divergence (cf. Tsairidis *et al.* (1996)). For the inequality $I_{qrc} \leq I_{rc}$, and for both measures, take the triplet (Z, Y, δ). Plainly, by the same principle, $I_{qrc} = I_{(\delta,Y)} \leq I_{(Z,Y,\delta)}$. On the other side, an application of the factorization theorem shows that the pair (Z, δ) is sufficient for (Z, Y, δ) and, thus, $I_{qrc} \leq I_{(Z,Y,\delta)} = I_{(Z,\delta)} = I_{rc}$. An alternative proof can be obtained by considering the additivity property of $I_{(\delta,Y)}$ and $I_{(Z,\delta)}$, and showing that $E\left(I_{Y|\delta}\right) \leq E\left(I_{Z|\delta}\right)$. For $\delta = 0$, $I_{Y|\delta=0} = I_{Z|\delta=0}$, since the random variables $Y \mid \delta = 0$ and $Z \mid \delta = 0$ are equal almost everywhere. The conditional distributions of $Y \mid \delta = 1$ and $Z \mid \delta = 1$ are given by $Cg(y)F(y)$ and $Cf(z)\bar{G}(z)$, respectively, where C is a normalizing constant (one can easily check that the normalizing constants are the same). The p.d.f. of $Y \mid X = x, \delta = 1$ is given by

$$\bar{g}_{Y|X=x,\delta=1}(y) = \begin{cases} \dfrac{g(y)}{\bar{G}(x)}, & y \geq x \\ 0, & \text{otherwise} \end{cases}$$

and so the density of $Y \mid \delta = 1$ is

$$\int_0^y \frac{g(y)}{\bar{G}(x)} Cf(x)\bar{G}(x)dx = Cg(y)F(y).$$

Thus, the inequality $I_{Y|\delta=1} \leq I_{Z|\delta=1}$ follows from the monotonicity or maximal information property of measures of information, which hold for both Fisher's and Csiszar's measures. A further analytical proof is available from the authors.

Remark 2.1. The second proof of Theorem 2.1 establishes a more general result: If X is a positive random variable with density $Cf(x)\bar{G}(x)$ and Y a random variable with density $Cg(y)F(y)$, where $F' = f$ and $G' = g$, then $I_Y \leq I_X$.

Remark 2.2. It is known that random censoring includes type I censoring by simply setting $Y = t_c$, where t_c is some (preassigned) fixed number called fixed censoring time (see, e.g., Miller (1981)). Setting $Y = t_c$, we have the recent result of Gertsbakh (1995), for the Fisher measure of information with θ univariate, in type I censored and quantal response data.

Ferentinos and Papaioannou (1981) considered byproducts of Fisher's information as parametric measures of information. Let λ_i^{rc} , and λ_i^{qrc} and λ_i^{δ}, $i = 1, 2, \ldots, k$, be the eigenvalues of the randomly censored and quantal randomly censored Fisher information matrices, respectively, and ι_{ii}^{rc}, and ι_{ii}^{qrc} and ι_{ii}^{δ}, $i = 1, 2, \ldots, k$, be the diagonal elements of the previous matrices, respectively. Then

$$D_{rc}(\theta) = \det\left[I_{rc}^F(\theta)\right] = \prod_{i=1}^{k} \lambda_i^{rc}, \qquad I_{rc}^*(\theta) = \mathrm{tr}[I_{rc}^F(\theta)] = \sum_{i=1}^{k} \iota_{ii}^{rc},$$

and

$$D_{qrc}(\theta) = \det[I_{qrc}^F(\theta)] = \prod_{i=1}^{k} \lambda_i^{qrc}, \qquad I_{qrc}^*(\theta) = \mathrm{tr}[I_{qrc}^F(\theta)] = \sum_{i=1}^{k} \iota_{ii}^{qrc},$$

$$D_{\delta}(\theta) = \det[I_{\delta}^F(\theta)] = \prod_{i=1}^{k} \lambda_i^{\delta}, \qquad I_{\delta}^*(\theta) = \mathrm{tr}[I_{\delta}^F(\theta)] = \sum_{i=1}^{k} \iota_{ii}^{\delta},$$

are the byproducts of the censored and quantal censored Fisher information, respectively. The following corollary shows that the byproducts of randomly censored Fisher information are greater than or equal to the byproducts of quantal randomly censored Fisher information.

COROLLARY 2.1. *For the byproducts of the randomly censored and quantal randomly censored Fisher information, we have*
(i) $D_{\delta}(\theta) \leq D_{qrc}(\theta) \leq D_{rc}(\theta)$, *and*
(ii) $I_{\delta}^*(\theta) \leq I_{qrc}^*(\theta) \leq I_{rc}^*(\theta)$.

PROOF. (i) Since $I_{qrc}^F(\theta)$, $I_{rc}^F(\theta) - I_{qrc}^F(\theta)$ and $I_{\delta}^F(\theta)$, $I_{qrc}^F(\theta) - I_{\delta}^F(\theta)$ are symmetric matrices, and $I_{rc}^F(\theta) - I_{qrc}^F(\theta)$ and $I_{qrc}^F(\theta) - I_{\delta}^F(\theta)$ are nonnegative definite, we have

$$\lambda_i^{rc} \geq \lambda_i^{qrc} \geq \lambda_i^{\delta}, \qquad i = 1, 2, \ldots, k,$$

(cf. Bellman (1970), p. 117, Theorem 3) where

$$\lambda_1^{rc} \geq \lambda_2^{rc} \geq \cdots \geq \lambda_k^{rc},$$
$$\lambda_1^{qrc} \geq \lambda_2^{qrc} \geq \cdots \geq \lambda_k^{qrc},$$

and

$$\lambda_1^{\delta} \geq \lambda_2^{\delta} \geq \cdots \geq \lambda_k^{\delta}.$$

So we have

$$D_{\delta}(\theta) \leq D_{qrc}(\theta) \leq D_{rc}(\theta).$$

(ii) Every diagonal element of the matrices $I_{rc}^F(\theta) - I_{qrc}^F(\theta)$ and $I_{qrc}^F(\theta) - I_{\delta}^F(\theta)$ are nonnegative. This implies

$$\mathrm{tr}[I_{rc}^F(\theta) - I_{qrc}^F(\theta)] \geq 0 \quad \text{and} \quad \mathrm{tr}\left[I_{qrc}^F(\theta) - I_{\delta}^F(\theta)\right] \geq 0,$$

that is,

$$I_{\delta}^*(\theta) \leq I_{qrc}^*(\theta) \leq I_{rc}^*(\theta).$$

Note that the $D_{\delta}(\theta) \leq D_{qrc}(\theta)$ and $I_{\delta}^*(\theta) \leq I_{qrc}^*(\theta)$ parts of the previous double inequalities follow from the monotonicity property.

3. Information in random censoring with random recording of uncensored observations

Let X be the lifetime response variable, and Y be the random inspection time, X and Y are assumed to be independent. If the item fails before inspection, i.e., if $X < Y$, then, either with probability p, $0 \le p \le 1$, the failure is immediately signalled, and the true lifetime $Z = X$ is recorded, or with probability $1 - p$, the failure is not signalled, and the item's true condition will be discovered only at the moment Y of inspection. If the lifetime $X > Y$, then the experiment terminates at the moment Y of inspection, and the censored lifetime $Z = Y$ is recorded.

This experimental situation can be thought as a mixture of quantal and complete randomly censored response. Elperin and Gertsbakh (1988) refer to this setup as random censoring with incomplete information. This is a randomly censored experiment, where the uncensored (true) observations are, either fully recorded $(X_i, \delta_i = 1)$ with probability p, $0 \le p \le 1$, or quantally (qualitatively) recorded $(Y_i, \delta_i = 1)$ with probability $1 - p$, and the censored observations are fully recorded $(Y_i, \delta_i = 0)$.

If $p = 0$, i.e., a failure is never signalled, we arrive at the random quantal response model (cf. Nelson (1982)). If $p = 1$, i.e., a failure is immediately displayed, we arrive at the random censoring model (see, e.g., Miller (1981) and Lawless (1982)).

Consider now an item i, which has lifetime X_i with c.d.f. $F(x;\theta)$, $\theta \in \Theta$. The item i is inspected at time Y_i (random), since the beginning of its operation. Let $G(y)$ be the c.d.f. of Y_i. Random variables X_i and Y_i are independent. Then, according to Elperin and Gertsbakh (1988), the likelihood associated with testing item i is

$$L(z_i;\theta) = \left[f(z_i;\theta)\bar{G}(z_i)p\right]^{\alpha\delta}\left[g(z_i)F(z_i;\theta)(1-p)\right]^{(1-\alpha)\delta}\left[g(z_i)\bar{F}(z_i;\theta)\right]^{1-\delta},$$

where z_i is the observed lifetime or the inspection time, and δ, α are indicator variables defined as follows:

$$\delta = \begin{cases} 1, & \text{if the lifetime is uncensored} \\ 0, & \text{if the lifetime is censored} \end{cases},$$

$$\alpha = \begin{cases} 1, & \text{if the failure is signalled} \\ 0, & \text{if the failure is not signalled} \end{cases}.$$

The likelihood associated with testing item i, given $Y_i = y_i$, is

$$L(z_i \mid y_i;\theta) = [f(z_i;\theta)p]^{\alpha\delta}[F(y_i;\theta)(1-p)]^{(1-\alpha)\delta}[\bar{F}(y_i;\theta)]^{1-\delta}. \tag{3.1}$$

By definition, the conditional Fisher information matrix for the experiment at hand (with a single observation) about θ, when $Y = y$, is

$$I^F_{(Z,\delta,\alpha)|Y=y}(\theta) = \left\| E\left(\frac{\partial \log L(z \mid y;\theta)}{\partial\theta_i}\frac{\partial \log L(z \mid y;\theta)}{\partial\theta_j}\right)\right\|_{k\times k}. \tag{3.2}$$

By relation (3.2), the Fisher information matrix is

$$I^F(\theta) \equiv I^F_{(Z,\delta,\alpha)}(\theta) = \int_0^\infty g(y) I^F_{(Z,\delta,\alpha)|Y=y}(\theta)dy.$$

In a similar manner, in view of (3.1),

$$p(z \mid y) = [f(z)p]^{\alpha\delta}[F(y)(1-p)]^{(1-\alpha)\delta}[\bar{F}(y)]^{1-\delta}, \tag{3.3}$$

is the conditional distribution for a randomly censored experiment with random recording of the uncensored observations, given $Y = y$. Then, the conditional φ-divergence measure and the φ-divergence measure, for discriminating between f_1 and f_2, are

$$I^C_{(Z,\delta,\alpha)|Y=y}(f_1, f_2) = \sum_{\alpha,\delta=0,1} E_{p_2}\left[\varphi\left(\frac{p_1(z \mid y)}{p_2(z \mid y)}\right)\right], \tag{3.4}$$

and

$$I^C(f_1, f_2) \equiv I^C_{(Z,\delta,\alpha)}(f_1, f_2) = \int_0^\infty g(y) I^C_{(Z,\delta,\alpha)|Y=y}(f_1, f_2)dy,$$

respectively.

The following theorem shows that measures of information for the experiment at hand are convex combinations, with weights p and $1-p$, of the randomly censored and the quantal randomly censored measures of information, respectively. The validity of (i) of the theorem, with θ univariate, is established by Elperin and Gertsbakh (1988).

THEOREM 3.1. *For the Fisher information matrix and the φ-divergence measure in a randomly censored experiment with random recording of the uncensored observations, we have*

(i) $I^F(\theta) = pI^F_{rc}(\theta) + (1-p)I^F_{qrc}(\theta)$, *and*
(ii) $I^C(f_1, f_2) = pI^C_{rc}(f_1, f_2) + (1-p)I^C_{qrc}(f_1, f_2)$.

PROOF. (i) For the sake of brevity we will use the following notation:

$$f(z;\theta) = f, \quad f_{(i)}(z;\theta) = f_{(i)}, \quad F(y;\theta) = F, \ F_{(i)}(y;\theta) = F_{(i)}.$$

By relations (3.1) and (3.2) we have

$$\begin{aligned} I^F_{(Z,\delta,\alpha)|Y=y}(\theta) = E\Bigg[&\alpha\left(\delta\frac{f_{(i)}f_{(j)}}{f^2} - \delta(1-\delta)\frac{f_{(i)}F_{(j)} + f_{(j)}F_{(i)}}{f\bar{F}} + (1-\delta)\frac{F_{(i)}F_{(j)}}{\bar{F}^2}\right) \\ &+(1-\alpha)\left(\delta\frac{F_{(i)}F_{(j)}}{F^2} - 2\delta(1-\delta)\frac{F_{(i)}F_{(j)}}{F\bar{F}} + (1-\delta)\frac{F_{(i)}F_{(j)}}{\bar{F}^2}\right)\Bigg]. \end{aligned} \tag{3.5}$$

But

$$\iota^{rc}_{ij} = E\left(\delta\frac{f_{(i)}f_{(j)}}{f^2} - \delta(1-\delta)\frac{f_{(i)}F_{(j)} + f_{(j)}F_{(i)}}{f\bar{F}} + (1-\delta)\frac{F_{(i)}F_{(j)}}{\bar{F}^2}\right), \tag{3.6}$$

and

$$\iota^{qrc}_{ij} = E\left(\delta\frac{F_{(i)}F_{(j)}}{F^2} - 2\delta(1-\delta)\frac{F_{(i)}F_{(j)}}{F\bar{F}} + (1-\delta)\frac{F_{(i)}F_{(j)}}{\bar{F}^2}\right), \tag{3.7}$$

where ι^{rc}_{ij} and ι^{qrc}_{ij}, $i,j = 1,2,\ldots,k$, are the (i,j) elements of the matrices

$$I^{F(rc)}_{(Z,\delta)|Y=y}(\theta) = \left\| \int_0^y \frac{f_{(i)}(x;\theta)f_{(j)}(x;\theta)}{f(x;\theta)}dx + \frac{F_{(i)}(y;\theta)F_{(j)}(y;\theta)}{\bar{F}(y;\theta)} \right\|_{k\times k},$$

and

$$I^{F(qrc)}_{(\delta,Y)|Y=y}(\theta) = \left\| \frac{F_{(i)}(y;\theta)F_{(j)}(y;\theta)}{F(y;\theta)\bar{F}(y;\theta)} \right\|_{k\times k},$$

respectively. Averaging with respect to α, one obtains from (3.5)–(3.7) that

$$I^F_{(Z,\delta,\alpha)|Y=y}(\theta) = P(\alpha=1)I^{F(rc)}_{(Z,\delta)|Y=y}(\theta) + P(\alpha=0)I^{F(qrc)}_{(\delta,Y)|Y=y}(\theta) = pI^{F(rc)}_{(Z,\delta)|Y=y}(\theta) + (1-p)I^{F(qrc)}_{(\delta,Y)|Y=y}(\theta). \tag{3.8}$$

By relation (3.8), averaging with respect to Y, we have

$$I^F(\theta) = pI^F_{rc}(\theta) + (1-p)I^F_{qrc}(\theta).$$

(ii) Similarly, by relations (3.3) and (3.4) we have

$$I^C_{(Z,\delta,\alpha)|Y=y}(f_1,f_2) = E_{p_2}\left\{\alpha\left[\delta\varphi\left(\frac{f_1(x)}{f_2(x)}\right) + (1-\delta)\varphi\left(\frac{\bar{F}_1(y)}{\bar{F}_2(y)}\right)\right] + (1-\alpha)\left[\delta\varphi\left(\frac{F_1(y)}{F_2(y)}\right) + (1-\delta)\varphi\left(\frac{\bar{F}_1(y)}{\bar{F}_2(y)}\right)\right]\right\}. \tag{3.9}$$

But

$$E_{p_2}\left[\delta\varphi\left(\frac{f_1(x)}{f_2(x)}\right) + (1-\delta)\varphi\left(\frac{\bar{F}_1(y)}{\bar{F}_2(y)}\right)\right] = I^{C(rc)}_{(Z,\delta)|Y=y}(f_1,f_2), \tag{3.10}$$

and

$$E_{p_2}\left[\delta\varphi\left(\frac{F_1(y)}{F_2(y)}\right) + (1-\delta)\varphi\left(\frac{\bar{F}_1(y)}{\bar{F}_2(y)}\right)\right] = I^{C(qrc)}_{(\delta,Y)|Y=y}(f_1,f_2). \tag{3.11}$$

Averaging with respect to α, one obtains from (3.9)–(3.11) that

$$I^C_{(Z,\delta,\alpha)|Y=y}(f_1,f_2) = P(\alpha=1)I^{C(rc)}_{(Z,\delta)|Y=y}(f_1,f_2) + P(\alpha=0)I^{C(qrc)}_{(\delta,Y)|Y=y}(f_1,f_2) = pI^{C(rc)}_{(Z,\delta)|Y=y}(f_1,f_2) + (1-p)I^{C(qrc)}_{(\delta,Y)|Y=y}(f_1,f_2). \tag{3.12}$$

By relation (3.12), averaging with respect to Y, we have

$$I^C(f_1,f_2) = pI^C_{rc}(f_1,f_2) + (1-p)I^C_{qrc}(f_1,f_2).$$

Let us now allow the probability of recording an observation to depend on the tested item i. Let this probability be p_i, and $q_i = 1-p_i$, $i=1,2,\ldots,n$. In this case it is easy to prove the following theorem.

THEOREM 3.2. *If $I^F_{sample}(\theta)$ and $I^C_{sample}(f_1,f_2)$ are the Fisher and Csiszar informations for the whole sample, respectively, then*

(i) $I^F_{sample}(\theta) = nI^F(\theta) = \left(\sum_{i=1}^n p_i\right)I^F_{rc}(\theta) + \left(\sum_{i=1}^n q_i\right)I^F_{qrc}(\theta)$, *and*

(ii) $I^C_{sample}(f_1,f_2) = nI^C(f_1,f_2) = \left(\sum_{i=1}^n p_i\right)I^C_{rc}(f_1,f_2) + \left(\sum_{i=1}^n q_i\right)I^C_{qrc}(f_1,f_2)$.

4. Information decomposition

Consider the randomly censored model, that is, suppose that $Z_i = \min(X_i, Y_i)$, $i=1,2,\ldots,n$, are i.i.d. random variables which represent the lifetimes (true or censored) of n items on test. Denote with f and F the p.d.f. and c.d.f. of X, respectively, and with

g the p.d.f. of Y. Let the censored experiment be noninformative, i.e., the distribution of Y be independent of the parameter θ. By $I_X^F(\theta)$ and $I_X^{KL}(f_1, f_2)$ we shall denote the classical measures of information for X.

Let $\bar{f}$ denote the p.d.f. of $X \mid Y = y, \delta = 0$. Then

$$\bar{f}_{X|Y=y,\delta=0}(x;\theta) = \begin{cases} \dfrac{f(x;\theta)}{\bar{F}(y;\theta)}, & x \geq y \\ 0, & \text{otherwise} \end{cases}. \tag{4.1}$$

Let also $I^{F(rc)}_{cens|Y=y}(\theta)$ denote the conditional (given that $Y = y$) Fisher information matrix on θ contained in a censored observation in random censoring. Then

$$I^{F(rc)}_{cens|Y=y}(\theta) = \left\| \int_y^\infty \bar{f}(x;\theta) \left(\frac{\partial}{\partial\theta_i} \log \bar{f}(x;\theta) \frac{\partial}{\partial\theta_j} \log \bar{f}(x;\theta) \right) dx \right\|_{k\times k}, \quad \theta \in \Theta. \tag{4.2}$$

In view of (4.1), the conditional (given that $Y = y$) Kullback-Leibler divergence between f_1 and f_2 based on a censored observation in random censoring, will be denoted by $I^{KL(rc)}_{cens|Y=y}(f_1, f_2)$, and is defined by

$$I^{KL(rc)}_{cens|Y=y}(f_1, f_2) = \int_y^\infty \bar{f}_1(x) \log \frac{\bar{f}_1(x)}{\bar{f}_2(x)} dx. \tag{4.3}$$

The following theorem gives the formula of the "total information" for the Fisher and Kullback-Leibler measures of information.

THEOREM 4.1. *If $I_X^F(\theta)$ and $I_X^{KL}(f_1, f_2)$ are the classical measures of information for X, and $I^{F(rc)}_{cens|Y=y}(\theta)$ and $I^{KL(rc)}_{cens|Y=y}(f_1, f_2)$ are the conditional (given that $Y = y$) measures of information for a censored observation in random censoring defined by* (4.2) *and* (4.3), *respectively, then the following relationships hold*

(i) $I_X^F(\theta) = I_{rc}^F(\theta) + E_g\left(\bar{F}(Y;\theta) I^{F(rc)}_{cens|Y}(\theta)\right)$, *and*

(ii) $I_X^{KL}(f_1, f_2) = I_{rc}^{KL}(f_1, f_2) + E_g\left(\bar{F}_1(Y) I^{KL(rc)}_{cens|Y}(f_1, f_2)\right)$.

These formulas will be referred to as "total information".

PROOF (i) By relation (4.2), and after some algebra, we have

$$I^{F(rc)}_{cens|Y=y}(\theta) = \frac{1}{\bar{F}(y;\theta)} \left\| \int_y^\infty \frac{f_{(i)}(x;\theta) f_{(j)}(x;\theta)}{f(x;\theta)} dx - \frac{F_{(i)}(y;\theta) F_{(j)}(y;\theta)}{\bar{F}(y;\theta)} \right\|_{k\times k}. \tag{4.4}$$

Multiplying the right hand side of (4.4) by $\bar{F}(y;\theta)$ and adding it to the right hand side of

$$I^{F(rc)}_{(Z,\delta)|Y=y}(\theta) = \left\| \int_0^y \frac{f_{(i)}(z;\theta) f_{(j)}(z;\theta)}{f(z;\theta)} dz + \frac{F_{(i)}(y;\theta) F_{(j)}(y;\theta)}{\bar{F}(y;\theta)} \right\|_{k\times k},$$

we obtain

$$I_X^F(\theta) = I^{F(rc)}_{(Z,\delta)|Y=y}(\theta) + \bar{F}(y;\theta) I^{F(rc)}_{cens|Y=y}(\theta). \tag{4.5}$$

Averaging (4.5) with respect to Y, we have that

$$I_X^F(\theta) = I_{rc}^F(\theta) + E_g(\bar{F}(Y;\theta) I^{F(rc)}_{cens|Y}(\theta)).$$

(ii) In a similar manner, by relation (4.3), we have that

$$I^{KL(rc)}_{cens|Y=y}(f_1,f_2) = \frac{1}{\bar{F}_1(y)}\left[\int_y^\infty f_1(x)\log\frac{f_1(x)}{f_2(x)}dx - \bar{F}_1(y)\log\frac{\bar{F}_1(y)}{\bar{F}_2(y)}\right]. \tag{4.6}$$

But the conditional randomly censored Kullback-Leibler divergence, when $Y = y$, is given by

$$I^{KL(rc)}_{(Z,\delta)|Y=y}(f_1,f_2) = \int_0^y f_1(z)\log\frac{f_1(z)}{f_2(z)}dz + \bar{F}_1(y)\log\frac{\bar{F}_1(y)}{\bar{F}_2(y)}. \tag{4.7}$$

Multiplying the right hand side of (4.6) by $\bar{F}_1(y)$ and adding it to the right hand side of (4.7), we obtain

$$I^{KL}_X(f_1,f_2) = I^{KL(rc)}_{(Z,\delta)|Y=y}(f_1,f_2) + \bar{F}_1(y)I^{KL(rc)}_{cens|Y=y}(f_1,f_2). \tag{4.8}$$

Averaging (4.8) with respect to Y, we have that

$$I^{KL}_X(f_1,f_2) = I^{KL}_{rc}(f_1,f_2) + E_g(\bar{F}_1(Y)I^{KL(rc)}_{cens|Y}(f_1,f_2)).$$

Now we establish analogous results for the Fisher and Kullback-Leibler measures of information, when instead of random censoring we have quantal random censoring or experiments with random recording of the uncensored observations. By Theorem 2.1(i) we have that

$$I^F_{rc}(\theta) - I^F_{qrc}(\theta) = \int_0^\infty g(y)F(y;\theta)B(\theta)dy, \qquad \theta\in\Theta,$$

where

$$B(\theta) = \left\| \int_0^y f^*(z;\theta)\left(\frac{\partial}{\partial\theta_i}\log f^*(z;\theta)\frac{\partial}{\partial\theta_j}\log f^*(z;\theta)\right)dz \right\|_{k\times k}, \qquad \theta\in\Theta,$$

with

$$f^*(z;\theta) = \frac{f(z;\theta)}{F(y;\theta)}, \qquad \theta\in\Theta, \tag{4.9}$$

the p.d.f. on $[0,y]$. By definition, the matrix $B(\theta)$ is the conditional Fisher information matrix on θ contained in a censored observation in quantal random censoring, given that $Y=y$, and, from now on, it will be denoted by $I^{F(qrc)}_{cens|Y=y}(\theta)$. So

$$I^F_{rc}(\theta) = I^F_{qrc}(\theta) + E_g(F(Y;\theta)I^{F(qrc)}_{cens|Y}(\theta)), \tag{4.10}$$

and taking into account Theorem 4.1(i), we have

$$I^F_X(\theta) = I^F_{qrc}(\theta) + E_g(F(Y;\theta)I^{F(qrc)}_{cens|Y}(\theta)) + E_g(\bar{F}(Y;\theta)I^{F(rc)}_{cens|Y}(\theta)), \tag{4.11}$$

which is the formula of the "total information" for Fisher type measure of information, associated with quantal randomly censored Fisher information.

Let us now consider a randomly censored experiment with random recording of uncensored observations. By Theorem 3.1(i) and relation (4.10), we have

$$I^F(\theta) = I^F_{qrc}(\theta) + p\cdot E_g(F(Y;\theta)I^{F(qrc)}_{cens|Y}(\theta)). \tag{4.12}$$

If we combine relations (4.11) and (4.12), we have

$$I_X^F(\theta) = I^F(\theta) + (1-p)\cdot E_g(F(Y;\theta)I_{cens|Y}^{F(qrc)}(\theta)) + E_g(\bar{F}(Y;\theta)I_{cens|Y}^{F(rc)}(\theta)),$$

which is the formula of the "total information" for Fisher type measure of information, associated with Fisher information in a randomly censored experiment with random recording of uncensored observations.

In view of (4.9), the conditional Kullback-Leibler randomly censored divergence between f_1 and f_2 based on a censored observation in quantal random censoring, given that $Y=y$, will be denoted by $I_{cens|Y=y}^{KL(qrc)}(f_1,f_2)$, and is defined by

$$(4.13)\qquad \begin{aligned} I_{cens|Y=y}^{KL(qrc)}(f_1,f_2) &= \int_0^y f_1^*(z)\log\frac{f_1^*(z)}{f_2^*(z)}dz \\ &= \frac{1}{F_1(y)}\left[\int_0^y f_1(z)\log\frac{f_1(z)}{f_2(z)}dz - F_1(y)\log\frac{F_1(y)}{F_2(y)}\right]. \end{aligned}$$

The conditional quantal randomly censored Kullback-Leibler divergence, when $Y=y$, is given by

$$(4.14)\qquad I_{(\delta,Y)|Y=y}^{KL(qrc)}(f_1,f_2) = F_1(y)\log\frac{F_1(y)}{F_2(y)} + \bar{F}_1(y)\log\frac{\bar{F}_1(y)}{\bar{F}_2(y)}.$$

Substracting (4.14) from (4.7), and multiplying both sides of (4.13) by $F_1(y)$, we obtain

$$(4.15)\qquad I_{(Z,\delta)|Y=y}^{KL(rc)}(f_1,f_2) = I_{(\delta,Y)|Y=y}^{KL(qrc)}(f_1,f_2) + F_1(y)I_{cens|Y=y}^{KL(qrc)}(f_1,f_2).$$

Averaging (4.15) with respect to Y, we have that

$$(4.16)\qquad I_{rc}^{KL}(f_1,f_2) = I_{qrc}^{KL}(f_1,f_2) + E_g(F_1(Y)I_{cens|Y}^{KL(qrc)}(f_1,f_2)).$$

Taking into account Theorem 4.1(*ii*), we have

$$(4.17)\qquad \begin{aligned} I_X^{KL}(f_1,f_2) &= I_{qrc}^{KL}(f_1,f_2) + E_g(F_1(Y)I_{cens|Y}^{KL(qrc)}(f_1,f_2)) \\ &\quad + E_g(\bar{F}_1(Y)I_{cens|Y}^{KL(rc)}(f_1,f_2)). \end{aligned}$$

For the randomly censored experiment with random recording of uncensored observations and in view of Theorem 3.1(ii), for $\varphi(u)=u\log u$ and relation (4.16), we have

$$(4.18)\qquad I^{KL}(f_1,f_2) = I_{qrc}^{KL}(f_1,f_2) + p\cdot E_g(F_1(Y)I_{cens|Y}^{KL(qrc)}(f_1,f_2)).$$

If we combine relations (4.17) and (4.18), we have

$$(4.19)\qquad \begin{aligned} I_X^{KL}(f_1,f_2) &= I^{KL}(f_1,f_2) + (1-p)\cdot E_g(F_1(Y)I_{cens|Y}^{KL(qrc)}(f_1,f_2)) \\ &\quad + E_g(\bar{F}_1(Y)I_{cens|Y}^{KL(rc)}(f_1,f_2)). \end{aligned}$$

Relations (4.17) and (4.19), give the formulas of the "total information" for Kullback-Leibler divergence, associated with quantal randomly censored Kullback-Leibler divergence and the Kullback-Leibler divergence in a randomly censored experiment with random recording of uncensored observations, respectively. The above lead to the following theorem.

THEOREM 4.2. *For the quantal randomly censored measures of information and the measures of information in a randomly censored experiment with random recording of uncensored observations, the following formulas for the total information hold*

(i) $I_X^F(\theta) = I_{qrc}^F(\theta) + E_g(F(Y;\theta)I_{cens|Y}^{F(qrc)}(\theta)) + E_g(\bar{F}(Y;\theta)I_{cens|Y}^{F(rc)}(\theta))$,

(ii) $I_X^F(\theta) = I^F(\theta) + (1-p) \cdot E_g(F(Y;\theta)I_{cens|Y}^{F(qrc)}(\theta)) + E_g(\bar{F}(Y;\theta)I_{cens|Y}^{F(rc)}(\theta))$,

(iii) $I_X^{KL}(f_1, f_2) = I_{qrc}^{KL}(f_1, f_2) + E_g(F_1(Y)I_{cens|Y}^{KL(qrc)}(f_1, f_2))$
$+ E_g(\bar{F}_1(Y)I_{cens|Y}^{KL(rc)}(f_1, f_2))$, *and*

(iv) $I_X^{KL}(f_1, f_2) = I^{KL}(f_1, f_2) + (1-p) \cdot E_g(F_1(Y)I_{cens|Y}^{KL(qrc)}(f_1, f_2))$
$+ E_g(\bar{F}_1(Y)I_{cens|Y}^{KL(rc)}(f_1, f_2))$.

The following theorem shows that, for both Fisher information matrix and Kullback-Leibler divergence, the property of being equivariant with respect to the expected value of the distribution function of lifetime, is equivalent to the *information lack-of-memory* property. The proof is straightforward.

THEOREM 4.3. (i) *The equivariance of Fisher information matrix with respect to the expected value of* $F(y;\theta)$, *i.e., the relation* $I_{rc}^F(\theta) = I_X^F(\theta) \cdot E_g\left(F(Y;\theta)\right)$, *is equivalent to the information lack-of-memory property*

$$I_{cens|Y=y}^{F(rc)}(\theta) = I_X^F(\theta).$$

(ii) *The equivariance of Kullback-Leibler divergence with respect to the expected value of* $F_1(y)$, *i.e., the relation* $I_{rc}^{KL}(f_1, f_2) = I_X^{KL}(f_1, f_2) \cdot E_g\left(F_1(Y)\right)$, *is equivalent to the information lack-of-memory property*

$$I_{cens|Y=y}^{KL(rc)}(f_1, f_2) = I_X^{KL}(f_1, f_2).$$

Remark 4.1. The first part of Theorems 4.1 and 4.3, with θ univariate, is given in Gertsbakh and Kagan (1999). It is easy to see that, relations (ii) and (iv) of Theorem 4.2, for $p = 0$, yield relations (i) and (iii) of the same theorem, respectively, and for $p = 1$, yield relations (i) and (ii) of Theorem 4.1, respectively.

Acknowledgements

The authors would like to thank the referees for their valuable comments and the updated references, which helped to considerably improve the presentation of the paper. One of the referees was very helpful in providing the alternative proofs of Theorem 2.1.

REFERENCES

Bellman, R. (1970). *Introduction to Matrix Analysis*, McGraw-Hill, New York.

Blackwell, D. (1951). Comparison of experiments, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 93–102, University of California Press, Berkeley.

Csiszar, I. (1963). Eine Informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizitat von Markoffschen Ketten, *A Mayar Tudomanyos Academia Matematikai Kutato Intezetenek Kozlemenyei*, **8**, 85–108.

Elperin, T. and Gertsbakh, I. (1988). Estimation in a random censoring model with incomplete information: Exponential lifetime distribution, *IEEE Transactions on Reliability*, **37**, 223–229.

Ferentinos, K. and Papaioannou, T. (1981). New parametric measures of information, *Information and Control*, **51**, 193–208.

Gertsbakh, I. (1995). On the Fisher information in type I censored and quantal response data, *Statist. Probab. Lett.*, **23**, 297–306.

Gertsbakh, I. and Kagan, A. (1999). Characterization of the Weibull distribution by properties of Fisher information under type I censoring, *Statist. Probab. Lett.*, **42**, 99–105.

Hollander, M., Proschan, F. and Sconing, J. (1987). Measuring information in right censored models, *Naval Res. Logist.*, **34** , 669–681.

Kullback, S. and Leibler, R. A. (1951). On information and sufficiency, *Ann. Math. Statist.*, **22**, 79–86.

Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*, Wiley, New York.

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

Miller, R. (1981). *Survival Analysis*, Wiley, New York.

Nelson, W. (1982). *Applied Life Data Analysis*, Wiley, New York.

Papaioannou, T. (1985). Measures of information,*Encyclopedia of Statistical Sciences* (eds. S. Kotz and N. L. Johnson) **5**, 391–397, Wiley, New York.

Pukelsheim, F. (1993). *Optimal Design of Experiments*, Wiley, New York.

Tsairidis, Ch., Ferentinos, K. and Papaioannou, T. (1996). Information and randon censoring, *Inform. Sci.*, **92**, 159–174.