Methodology and Computing in Applied Probability (2025) 27:27
https://doi.org/10.1007/s11009-025-10153-x

A Conditioned Kullback-Leibler Divergence Measure through Compensator Processes and its Relationship to Cumulative Residual Inaccuracy Measure with Applications

Vanderlei da Costa Bueno[1] · **Narayanaswamy Balakrishnan**[2]

Received: 26 April 2024 / Revised: 13 February 2025 / Accepted: 8 March 2025

Abstract

Kullback-Leibler divergence measure between two random variables is quite useful in many contexts and has received considerable attention in numerous fields including statistics, physics, probability, and reliability theory. A cumulative Kullback-Leibler divergence measure has been proposed recently as a suitable extension of this measure upon replacing density functions by cumulative distribution functions. In this paper, we study a dynamic version of it by using a point process martingale approach conditioned on an observed past. Interestingly, this concept is identical to cumulative residual inaccuracy measure introduced by (Bueno and Balakrishnan (Probab Eng Sci 36:294-319, 2022). We also extend the concept of relative cumulative residual information generating measure to a conditional one and get Kullback-Leibler divergence measure through it. We further extend the new versions to non-explosive univariate point processes. In particular, we apply the conditioned Kullback-Leibler divergence to compare measures between two non-explosive point processes. Several applications of the established results are presented, including to a general repair process, minimal repair point process, coherent systems, Markov-modulated Poisson processes and Markov chains.

Keywords Conditioned Kullback-Leibler divergence measure · Conditional relative cumulative residual information generating measure · Cumulative residual inaccuracy measure · Point process martingale · Stochastic inequalities · General repair processes

Mathematics Subject Classification (2010) Primary: 90B25 · Secondary: 60G44

✉ Vanderlei da Costa Bueno
bueno@ime.usp.br

Narayanaswamy Balakrishnan
bala@mcmaster.ca

1 Institute of Mathematics and Statistics, São Paulo University, São Paulo, Brazil

2 Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario L8S 4LS, Canada

Published online: 25 March 2025

1 Introduction

Consider two positive random variables T and S defined in a probability space $(\Omega, \Im, P)$, with distribution functions $F(t)$ and $G(t)$ and continuous density functions $f(t)$ and $g(t)$, respectively. Then Shannon entropy, Shannon (1948), a measure of uncertainty, is given by

$$H(f) = -\int_0^\infty f(t) \ln f(t) dt$$

where ln is the natural logarithm, the base e logarithm function.

An important inaccuracy measure between the uncertainty of two positive and absolutely continuous random variables, S and T, is Kerridge inaccuracy measure, Kerridge (1961), given by

$$H(S, T) = E[-\ln g(T)] = -\int_0^\infty (\ln g(x)) f(x) dx.$$

In the case when S and T are identically distributed, the Kerridge inaccuracy measure reduces to the well-known Shannon entropy measure.

Rao et al. (2004); Rao (2005) provided an extension of the Shannon entropy, called the cumulative residual entropy of T, $CRE(T)$, by using survival functions instead of probability density functions, as

$$CRE(T) = -\int_{-\infty}^\infty \overline{F}(t) \ln \overline{F}(t) dt,$$

where $\overline{F} = 1 - F$ is the survival function.

Kerridge measure of inaccuracy has also been extended in a similar way by Kumar and Taneja (2012, 2015).

The Kumar and Taneja's cumulative residual inaccuracy measure between S and T is defined as

$$\varepsilon(S, T) = -\int_0^\infty \overline{F}(t) \ln \overline{G}(t) dt = E[\int_0^T \Lambda_S(s) ds],$$

where $\overline{F} = 1 - F$ and $\overline{G} = 1 - G$ are the reliability functions of T and S, respectively, and $\Lambda_S(t) = -\log \overline{G}(t)$ is the cumulative hazard function of S. We take, as convention, $0 \log 0 = 0$.

Indeed, $\varepsilon(S, T)$ represents the information content when using $\overline{G}(t)$, the survival function asserted by the experimenter, due to missing/incorrect information, instead of the true survival function $\overline{F}(t)$.

A main measure of information, named Kullback-Leibler (KL) information, is defined as in Kullback and Leibler (1951)

$$I(S, T) = \int_{-\infty}^\infty g(t) \ln \left(\frac{g(t)}{f(t)} \right) dt,$$

is a measure of information discrepancy between T and another random variable S. Here $g(t)$ can be viewed as a reference distribution.

The likelihood ratio in the above expression, may be considered as the information resulting from the observation $S = t$ for discrimination in favor of S against T. Therefore, the expression is then a mean discrimination.

As $t - 1 \geq \ln t, \quad \forall t > 0$, it is easy to prove that the information $I(S, T)$ have the nonnegative property and the equality $I(S, T) = 0$ holds if, and only if, $f(t) = g(t), \quad \forall t$.

However, KL information is not symmetric and cannot be a measure.

Inspired by $\varepsilon(S,T)$ of Kumar and Taneja (2015); Park et al. (2012) proposed a simple extension for positive random variables

$$\int_0^\infty \overline{G}(t)\ln(\frac{\overline{G}(t)}{\overline{F}(t)})dt.$$

However

$$\int_0^\infty \overline{G}(t)\ln(\frac{\overline{G}(t)}{\overline{F}(t)})dt = -\int_0^\infty \overline{G}(t)\ln(\frac{\overline{F}(t)}{\overline{G}(t)})dt$$

$$\geq \int_0^\infty \overline{G}(t)(1-\frac{\overline{F}(t)}{\overline{G}(t)})dt = \int_0^\infty \overline{G}(t)dt - \int_0^\infty \overline{F}(t))dt = E[S]-E[T]$$

which does not satisfy the nonnegativity and characterization properties.

To bypass this problem, Barrapour and Rad (2012) has suggested an extension of KL information using survival distribution function which can be called the residual KL information ($CRKL$) given by

$$CRKL(S,T) = \int_0^\infty \overline{G}(t)\ln(\frac{\overline{G}(t)}{\overline{F}(t)})dt - (E[S]-E[T]).$$

Another extension of this measure, using distribution functions was considered by Park et al. (2012) , which can be called cumulative KL information, (CKL), is

$$CKL(S,T) = \int_0^\infty G(t)\ln\frac{G(t)}{F(t)}dt + (E[S]-E[T])$$

which also satisfies the nonnegativity and characterization properties, but is not symmetric. Di Crescenzo and Longobardi (2015) analyse some properties and applications of this particular measure.

To obtain the symmetric property, Kullback and Leibler (1951) define a (KL) divergence measure by

$$J(S,T) = I(S,T) + I(T,S) = \int_{-\infty}^\infty g(t)\ln\left(\frac{g(t)}{f(t)}\right)dt + \int_{-\infty}^\infty f(t)\ln\left(\frac{f(t)}{g(t)}\right)dt.$$

The above expression, introduced by Jeffreys (1946, 1948), defines the divergence between T and S as a measure of the difficulty of discriminating between them.

We observe the following properties of $J(S,T)$

I) Since that $t-1 \geq \ln t, \quad \forall t > 0$ we have

$$I(S,T) = \int_{-\infty}^\infty g(t)\ln\left(\frac{g(t)}{f(t)}\right)dt = -\int_{-\infty}^\infty g(t)\ln\left(\frac{f(t)}{g(t)}\right)dt$$

$$\geq \int_0^\infty g(t)(1-\frac{f(t)}{g(t)})dt = \int_0^\infty g(t)dt - \int_0^\infty f(t))dt = 0.$$

Therefore $J(S,T) = I(S,T) + I(T,S)$ is positive.

II) $J(S,T) = I(S,T) + I(T,S) = 0$ if and only if $f(t) = g(t) \quad \forall t > 0$. Otherwise $|(f(t)-g(t))\ln\left(\frac{f(t)}{g(t)}\right)| > 0$ and $J(S,T) \neq 0$.

III) $J(S,T) = I(S,T) + I(T,S)$ is symmetric.

IV) For any random variables S, T and U defined in $(\Omega, \Im, P)$, the triangle inequality is defined as

$$J(S,T) \leq J(S,U) + J(U,T).$$

When the random variables are not equal almost everywhere and the properties of non-negativity (I), identity of indiscernibles (II), symmetry (III) and triangle inequality (IV)worth we say the $J(S,T)$ is a metric in the L^1-space of random variables.

However, if S, T and U are independent random variables with exponential distributions with parameters α, β and γ, respectively we have

$$J(S,T) = \frac{(\beta-\alpha)^2}{\alpha.\beta}$$

and the triangle inequality becames

$$\frac{(\beta-\alpha)^2}{\alpha.\beta} \le \frac{(\alpha-\gamma)^2}{\alpha.\gamma}\frac{(\gamma-\beta)^2}{\gamma.\beta}$$

If we replace $\alpha = 1$, $\beta = 3$ and $\gamma = 2$ we have $\frac{4}{3} \le \frac{1}{6} + \frac{1}{2} = \frac{4}{6}$ is not true.

Therefore, as Kullback (1968) described, $J(S,T)$ is symmetric and has all the properties of a metric, except the triangular inequality property. The information measures $I(S,T)$ and $I(T,S)$ may, in this respect, be considered as directed distances.

In information theory generating functions have been defined for probability density functions (PDFs) to determine information quantities such as Shannon information, informational energy, extropy and Kullback-Leibler information. Golomb (1966) proposed information generating function of a PDF f, whose derivatives, evaluated at 1, yeld some statistical information measures for the probability distributions. To this end, let T be an absolutely continuous random variable defined in a complete probability space $(\Omega, \Im, P)$, with cumulative distribution function F and probability density function f. Then, the information generating (IG) function of T , for any $\alpha > 0$, is defined as

$$G_\alpha(T) = \int_{-\infty}^{\infty} f^\alpha(t)dt,$$

provided the integral exists. Golomb (1966) then showed the following properties of $G_\alpha(T)$:

$$G_1(T) = 1 \ \ and \ \ \frac{\mathrm{d}}{\mathrm{d}\alpha}G_\alpha(T)|_{\alpha=1} = -H(T),$$

where $\frac{\mathrm{d}}{\mathrm{d}\alpha}G_\alpha(T)$ is the first derivative of $G_\alpha(T)$, with respect to α and $H(T)$ is the Shanoon differential entropy. In particular, when $\alpha = 2$, the IG measure becomes $\int_{\infty}^{\infty} f^2(t)dt$, known as informational energy (IE) function.

Relative entropy (based in Golomb's information function) has been defined by Guiasu and Reischer (1985); its first derivative at 1 yelds Kullback-Leibler divergence. Let T and S be two absolutely continuous random variables defined on a complete probability space $(\Omega, \Im, P)$, with probability density functions f and g, respectively. Then, the relative generating (RIG) function, for any $\alpha > 0$, is defined as

$$R_\alpha(T,S) = \int_{-\infty}^{\infty} f^\alpha(t)g^{1-\alpha}(t)dt,$$

provided the integral exists. The Kullback and Leibler information is then obtained as

$$KL(T,S) = \frac{d}{d\alpha}R_\alpha(T,S)|_{\alpha=1} = \int_{-\infty}^{\infty} f(t)\ln\left(\frac{f(t)}{g(t)}\right)dt.$$

Recently there is a deep interest in proposing new versions of information generating functions. Clark (2020) propose a measure applicable to point processes. Kharazmi and

Balakrishnan (2021) propose Jensen-information generating function, whose derivatives generate the well-known Jensen-Shannon, Jensen-Taneja and Jensen-Extropy information measures. Furthermore develope some results for $G_\alpha(T)$ for the residual lifetime distribution. Kharazmi and Balakrishnan (2023) proposed a cumulative residual information generating (CRIG) and relative cumulative residual information generating (RCRIG) functions based on survival functions.

For a non-negative random variable T with an absolutely continuous survival function $\overline{F}$, the cumulative residual information generating (CRIG) measure of T, for any $\alpha > 0$, is defined as

$$CIG_\alpha(T) = \int_0^\infty \overline{F}^\alpha(t)dt.$$

Let T and S be two absolutely continuous random variables defined on a complete probability space $(\Omega, \Im, P)$, with survival functions $\overline{F}$ and $\overline{G}$, respectively. Then, the relative cumulative residual information generating (RCRIG) measure between T and S, for any $\alpha > 0$, is defined as

$$RCRIG_\alpha(T, S) = \int_0^\infty \overline{F}^\alpha(t)\overline{G}^{1-\alpha}(t)dt$$

$$= \int_0^\infty \overline{F}(t)e^{(\alpha-1)\ln \overline{F}(t)}e^{(1-\alpha)\ln \overline{G}(t)}dt$$

$$= E\left[\int_0^T (\frac{e^{\ln \overline{F}(t)}}{e^{\ln \overline{G}(t)}})^{(\alpha-1)}dt\right].$$

Capaldo et al. (2024) introduce and study the cumulative generating function, which provides a unifying mathematical tool suitable to deal with classical entropies based on the cumulative distributions and on survival functions. In particular, extends (Kharazmi and Balakrishnan 2023) measure and the Gini mean semi-difference.

To explore information generating functions linked to maximum and minuimum ranked sets, as well as record values and their properties, refer to the works of Zamani et al. (2022).

We think that there is a gap in analysing relative cumulative residual information generating and its generated measures, such as KL measure, in the case of variables, or point processes, stochastically dependents. Martingale theory provides an interesting and elegant approach to work stochastically dependence and time dependence.

The rest of this paper consists of three sections. In Section 2, we provide motivation and purpose of this paper. In Section 3.1, of Section 3, we propose a relative conditional cumulative residual information generating measure (CRCRIG) and provide an interpretation for it. This information generating measure will be used to measure the closeness between two survival functions as well as to generate a conditional cumulative Kullback-Leibler (CKL) divergence measure. In Section 3.2, an extension of information generating measure to univariate non-explosive point processe is provided. In Section 4, we give some preliminary notation and a theorem about stochastic inequalities between non-explosive point processes through their compensator processes, mainly from Kwiecinski and Szekli (1991), to compare CKL divergence measures between two non-explosive point processes. We also describe several applications of the obtained results to general repair processes, minimal repair point processes, coherent systems, Markov-modulated Poisson processes and Markov chains. We present concluding remarks in Section 6.

2 Motivation and Purpose of this Work

Here, we extend the Kullback-Leibler (KL) divergence to a conditioned measure through compensator transformation.

As in Bueno and Balakrishnan (2022), we assume observing two component lifetimes T and S, which are finite positive absolutely continuous random variables defined on a complete probability space $(\Omega, \Im, P)$, with $P(S \neq T) = 1$, through the family of sub σ-algebras $(\Im_t)_{t\geq 0}$ of $\Im$, where

$$\Im_t = \sigma\{1_{\{S>s\}}, 1_{\{T>s\}}, 0 \leq s < t\}$$

satisfies Dellacherie's conditions of right continuity and completeness.

In this paper, in our general setup, for simplifying the notation, we assume that relations such as $\subset, =, \leq, <, \neq$ between random variables and measurable sets always hold with probability 1, which means that the term P-a.s. can be suppressed.

We now assume that S and T are totally inaccessible $\Im_t$-stopping times. An extended and positive random variable τ is a $\Im_t$-stopping time if, and only if, $\{\tau \leq t\} \in \Im_t$, for all $t \geq 0$; a $\Im_t$-stopping time τ is said to be predictable if an increasing sequence $(\tau_n)_{n\geq 0}$ of $\Im_t$-stopping times, $\tau_n < \tau$, exists such that $\lim_{n\to\infty} \tau_n = \tau$; a $\Im_t$-stopping time τ is totally inaccessible if $P(\tau = \sigma < \infty) = 0$ for all predictable $\Im_t$-stopping times σ. In this way, absolutely continuous lifetimes are thought of as totally inaccessible $\Im_t$-stopping times. For a mathematical basis of stochastic processes applied to reliability theory, one may refer to the books by Aven and Jensen (1999) and Bremaud (1981).

With respect to $(\Im_t)_{t\geq 0}$ and using Doob-Meyer decomposition, we consider the predictable compensator processes $(A_t)_{t\geq 0}$ and $(B_t)_{t\geq 0}$ such that $1_{\{T\leq t\}} - A_t$ and $1_{\{S\leq t\}} - B_t$ are 0 means uniformly integrable $\Im_t$-martingales. From the total inaccessibility of S and T, A_t and B_t are continuous.

It is well known, see (Jacod 1975), that there exists a bijective relationship between the space of all distributions functions and the $\Im_t$-compensator processes space characterized by the so called Doléans exponential equation

$$\overline{F}(t|\Im_t) = e^{-A_t^c} \pi_{s\leq t}(1 + \Delta A_s)$$

where A_t^c is the continuous part of A_t and $\Delta A_t = A_t - A_t^c$ is its discrete part.

Also, under the assumptions that T is a totally inaccessible $\Im_t$-stopping time, the $\Im_t$-compensators are continuous and from Arjas and Yashin (1988) we conclude that

$$A_t = -\ln \overline{F}(t \wedge T),$$

where $\wedge$ is the minimum operator, that is, $t \wedge T = \min\{t, T\}$.

We observe that, if A_t is a determinist increasing function of t (except that it is stopped when T occurs) we could say that T is dynamically independent of everything else in the history $\Im_t$. This is a generalization of the case of independent components: if the component lifetime T has a continuous compensator which is deterministic, then the lifetimes of the other variables have no causal effect in T. However, other components may well dependent casualy on T, so that the components need not be statistically independent. Furthermore, as in Norros (1986), the compensator at its final points are random variables independent and identically distributed standard exponential.

Therefore, by equivalence results between distribution functions and compensator processes, identifying $\bigwedge_S(t)$ and B_t, in the set $\{T \wedge S > t\}$, we obtain

$$E\left[\int_0^S A_t dt\right] = E\left[\int_0^S \left(\int_0^t dA_s\right) dt\right] = E\left[\int_0^S \left(\int_s^S dt\right) dA_s\right] = E\left[\int_0^S (S-s) dA_s\right].$$

As $\psi(s) = S - s$ is a left-continuous function, it is a $\Im_s$-predictable process and, as in Theorem T6 from Bremaud (1981), the integration of $\psi(s)$ with respect to the bounded variations martingales, $1_{\{T \leq t\}} - A_t$

$$M_s = \int_0^S (S-s) d(1_{\{T \leq s\}} - A_s)$$

is a mean 0 $\Im_s$-martingale. We then have

$$E\left[\int_0^S \bigwedge_T(t) dt\right] = E\left[\int_0^S (S-s) dA_s\right] = E\left[\int_0^S (S-s) d1_{\{T \leq t\}}\right]$$
$$= E\left[1_{\{T \leq S\}}(S-T)\right] = E\left[1_{\{T \leq S\}}|S-T|\right].$$

Also, with the same arguments, we have

$$E\left[\int_0^T \bigwedge_T(t) dt\right] = E\left[\int_0^T (T-s) dA_s\right] = E\left[\int_0^T (T-s) d1_{\{T \leq t\}}\right]$$
$$= E\left[1_{\{T \leq T\}}(T-T)\right] = 0,$$
$$E\left[\int_0^T \bigwedge_S(t) dt\right] = E\left[1_{\{S \leq T\}}|T-S|\right]$$

and

$$E\left[\int_0^S \bigwedge_S(t) dt\right] = E\left[\int_0^S (S-s) dB_s\right] = 0.$$

Furthermore, we consider the following extension of Kullback-Leibler (KL) divergence

Definition 2.1 Let S and T be continuous positive random variables defined on a complete probability space $(\Omega, \Im, P)$. Then, the **conditioned** Kullback-Leibler (CKL) divergence is defined as

$$J^C(S,T) = E\left[\int_0^T B_s ds\right] + E\left[\int_0^S A_s ds\right]$$
$$= E[1_{\{S \leq T\}}|T-S|] + E[1_{\{T \leq S\}}|S-T|] = E[|T-S|],$$

where $1_{\{T \leq t\}} - A_t$ and $1_{\{S \leq t\}} - B_t$ are 0 means uniformly integrable $\Im_t$-martingales.

Thus, $J^C(S,T)$ can be seen as a dispersion measure when using a lifetime S asserted by the experimenter's information of the true lifetime T. When the random variables are not equal almost everywhere, $J^C(T,S)$ is a metric in the L^1-space of random variables. Interestingly, this concept is identical to the cumulative residual inaccuracy measure from Bueno and Balakrishnan [5] which produces several results and applications.

If T and S satisfy the proportional risk hazard process, then the dynamic conditioned divergence measure, $J_t^C(T,S) < \infty$, uniquely determines the distribution function of T.

Bueno and Balakrishnan [6] extended $J_t^C(T,S)$, denoted in that paper by $CRI(T,S)$, for non-explosive point processes, as follows

Suppose $\mathbf{T} = (T_n)_{n\geq 1}$ and $\mathbf{S} = (S_n)_{n\geq 1}$ are two univariate non-explosive point processes observed at

$$\Im_t = \sigma\{1_{\{S_i>s\}}, 1_{\{Tj>s\}}, 0 \leq s < t, i, j \in \mathbf{N}\}$$

satisfying Dellacherie's conditions of right continuity and completeness and where T_i, $i \in \mathbf{N}$, and S_j, $j \in \mathbf{N}$, are totally unaccessible $\Im_t$-stopping times, with $P(T_i \neq S_j) = 1$, $\forall i, j \in \mathbf{N}$, and we consider the superposition of $\mathbf{T}$ and $\mathbf{S}$.

Definition 2.2 The superposition of two univariate point processes $\mathbf{T} = (T_n)_{n\geq 0}$ and $\mathbf{S} = (S_n)_{n\geq 0}$, defined on a complete probability space $(\Omega, \Im, P)$ with compensator processes $(A_t)_{t\geq 0}$ and $(B_t)_{t\geq 0}$, respectively, is the marked point process $(V_n, U_n)_{n\geq 1}$, where $\mathbf{V} = (V_n)_{n\geq 0}$ is a univariate point process, and $\mathbf{U} = (U_n)_{n\geq 0}$, the indicator process, is a sequence of random variables taking values in a measurable space $(\{0, 1\}, \sigma(\{0, 1\})$, resulting from pooling together the time points of events occurring in each of the separate point processes. Here, 0 stands for an occurrence of the process $\mathbf{T}$, $U_n = T_k$, for some k, in which case $V_n = \max_{1\leq j\leq n}\{(1 - U_j).V_j\}$, and 1 stands for an occurrence of the process $\mathbf{S}$, $U_n = S_j$ for some j, in which case $V_n = \max_{1\leq j\leq n}\{U_j.V_j\}$.

Definition 2.3 Let $\mathbf{T} = (T_n)_{n\geq 0}$ and $\mathbf{S} = (S_n)_{n\geq 0}$ be point processes with $\Im_t^V$-compensator processes $(A_t)_{t\geq 0}$ and $(B_t)_{t\geq 0}$, respectively, defined on a complete probability space $(\Omega, \Im, P)$. Let $(V_n, U_n)_{n\geq 1}$ be their superposition process. Then, the cumulative residual inaccuracy measure, at time t, between $\mathbf{T}$ and $\mathbf{S}$ is given by

$$J_t^C(\mathbf{T}, \mathbf{S}) = E\left[\int_0^t A_s ds + \int_0^t B_s ds\right]$$

$$= E\left[\int_0^t \sum_{n=1}^{\infty} \int_0^s 1_{\{U_n=0\}} dA_u^n ds + \int_0^t \sum_{n=1}^{\infty} \int_0^t 1_{\{U_n=1\}} dB_u^n ds\right].$$

Remark 2.4 Bueno and Balakrishnan (2024) remarked that when the random variables are not equal almost everywhere, the quantity $J(\mathbf{T}, \mathbf{S}) = E[\Sigma_{k=1}^{\infty}|V_k - V_{k-1}|]$ can be seen as a dispersion measure in the L^1-space of random variable sequences when using the point process $\mathbf{S}$, asserted by the experimenter's information of the true point process $\mathbf{T}$.

3 Information Measure Between Two Lifetimes

3.1 Conditional Relative Cumulative Residual Information Generating Functions

Suppose we observe two component lifetimes, T and S, which are finite positive random variables defined on a complete probability space $(\Omega, \Im, P)$, with $P(S \neq T) = 1$, through the family of sub σ-algebras $(\Im_t)_{t\geq 0}$ of $\Im$, where

$$\Im_t = \sigma\{1_{\{S>s\}}, 1_{\{T>s\}}, 0 \leq s < t\}$$

satisfies Dellacherie's conditions of right continuity and completeness.

With respect to $(\Im_t)_{t\geq 0}$ and using Doob-Meyer decomposition, we consider the predictable compensator processes $(A_t)_{t\geq 0}$ and $(B_t)_{t\geq 0}$ such that $1_{\{T\leq t\}} - A_t$ and $1_{\{S\leq t\}} - B_t$ are 0 means $\Im_t$-martingales. From the total inaccessibility of S and T, A_t and B_t are continuous.

The compensator process is expressed in terms of conditional probabilities, given the available information, and it generalizes the classical notion of hazard. Intuitively, it corresponds to producing whether the failure is going to occur now, on the basis of all observations available up to, but not including the present.

It then follows, by well-known equivalence results between distribution functions and compensator processes, that $A_t = -\ln \overline{F}(t \wedge T)$ and $B_t = -\ln \overline{G}(t \wedge S)$; see (Arjas and Yashin 1988). Identifying $-\ln \overline{G}(t)$ and B_t, and $-\ln \overline{F}(t)$ and A_t in the set $\{S \wedge T > t\}$, we present the following definition.

Definition 3.1.1 If S and T are continuous positive random variables defined on a complete probability space $(\Omega, \Im, P)$, the conditional relative cumulative residual information generating functions is

$$CRCRIG_\alpha(T,S) = CRIG_\alpha(T,S) + CRIG_\alpha(S,T) = E\left[\int_0^T e^{(\alpha-1)(B_t-A_t)}dt\right] + E\left[\int_0^S e^{(\alpha-1)(A_t-B_t)}dt\right].$$

provided the integral exists.

Remark 3.1.2 To clarify the conditions under which the integral $\int_0^T e^{(1-\alpha)(B_t-A_t)}dt$ we review the definitions of the $\Im_t$-compensators A_t and B_t

$$A_t = -\ln \overline{F}(t),\ \text{if } t < T \text{ and } A_t = A_T = -\ln \overline{F}(T)\ \textit{if}\ t \geq T.$$

$$B_t = -\ln \overline{G}(t),\ \text{if } t < S \text{ and } B_t = B_S = -\ln \overline{G}(S)\ \textit{if}\ t \geq S.$$

Furthermore A_T and B_S are independent and identically distributed standard exponential random variables, see (Norros 1986).

As A_t and B_t are increasing processes we have

$$E[\int_0^T (\frac{e^{(1-\alpha)B_t}}{e^{(1-\alpha)A_t}})dt] \leq E[\int_0^\infty 1_{\{T>t\}}(\frac{e^{(1-\alpha)B_T}}{e^{(1-\alpha)A_T}})dt] \leq E[\int_0^\infty 1_{\{T>t\}}(\frac{e^{(1-\alpha)B_S}}{e^{(1-\alpha)A_T}})dt].$$

However, if $1 < \alpha < 2$ we have

$$E[\int_0^\infty 1_{\{T>t\}}(\frac{e^{(1-\alpha)B_S}}{e^{(1-\alpha)A_T}})dt] = E\{E[\int_0^\infty 1_{\{T>t\}}(\frac{e^{(1-\alpha)B_S}}{e^{(1-\alpha)A_T}})dt|A_T]\} = E\{\int_0^\infty \frac{1_{\{A_T>A_t\}}}{e^{(1-\alpha)A_T}} E[e^{(1-\alpha)B_S}|A_T]dt\}$$

$$= (\frac{1}{\alpha})E\{\int_0^\infty \frac{1_{\{A_T>A_t\}}}{e^{(1-\alpha)A_T}}dt\} = (\frac{1}{\alpha})E\{\int_0^\infty 1_{\{A_T>A_t\}}e^{(\alpha-1)A_T}dt\}$$

$$= (\frac{1}{\alpha})(\frac{1}{\alpha-2})\int_0^\infty \overline{F}(t)dt = (\frac{1}{\alpha})(\frac{1}{\alpha-2})E[T].$$

The last equality follows from the equivalence $\{T > t\}$ and $\{A_T > A_t\}$ Therefore the conditions under which $E[T] < \infty$ and $E[S] < \infty$ are sufficient for integrability on Definition 3.1.1

At this point, we can calculate the fist derivative of $CRCIG_\alpha(T,S)$ with respect to α and evaluate at $\alpha = 1$:

$$\frac{\delta}{\delta\alpha}CRCRIG_\alpha(T,S) = E\left[\int_0^T (B_t - A_t)e^{(\alpha-1)(B_t-A_t)}dt\right] + E\left[\int_0^S (A_t - B_t)e^{(\alpha-1)(A_t-B_t)}dt\right]$$

and

$$\frac{\delta}{\delta\alpha}CRCRIG_\alpha(T,S)|_{\alpha=1} = E\left[\int_0^T (B_t - A_t)dt\right] + E\left[\int_0^S (A_t - B_t)dt\right] = E\left[\int_0^T B_t dt\right] + E\left[\int_0^S A_t dt\right].$$

Then, we can recall Definition 2.1 and verify that

$$\frac{d}{d\alpha} CRCRIG_\alpha(T, S)|_{\alpha=1} = E[|T - S|] = J^C(S, T),$$

that is, the first derivative of the conditional relative cumulative residual information generating measure, at $\alpha = 1$, is equal to KL divergence measure.

Interestingly, this concept is identical to the cumulative residual inaccuracy measure $CRI(S, T)$, of Bueno and Balakrishnan (2024). As such, the several results and properties of it have been proved by these authors.

Thus, $J^C(S, T)$ can be seen as a dispersion measure when using a lifetime S asserted by the experimenter's information of the true lifetime T. When the random variables are not equal almost everywhere, $J^C(T : S)$ is a metric in the L^1-space of random variables.

In the following we present an interpretation for $CRCRIG_\alpha(T, S)$.

Theorem 3.1.3 *If S and T are continuous positive random variables defined on a complete probability space $(\Omega, \Im, P)$, the conditional relative cumulative residual information generating function is*

$$CRCRIG_\alpha(T, S) = E[T] + E[S] + E[|T - S|e^{(\alpha-1)(|B_{T\wedge S} - A_{T\wedge S})|}].$$

Proof We have

$$CRIG_\alpha(T, S) = E\left[\int_0^T \left(\frac{e^{-A_t}}{e^{-B_t}}\right)^{(\alpha-1)} dt\right] = E\left[\int_0^T e^{(\alpha-1)(B_t - A_t)} dt\right].$$

However,

$$e^{(\alpha-1)(B_t - A_t)} = 1 + (\alpha - 1)\int_0^t e^{(\alpha-1)(B_s - A_s)} d(B_s - A_s)$$

and so

$$\begin{aligned}
CRIG_\alpha(T, S) &= E\left[\int_0^T (1 + (\alpha - 1)\int_0^t e^{(\alpha-1)(B_s - A_s)} d(B_s - A_s))dt\right] \\
&= E[T] + E\left[\int_0^T (\alpha - 1)\int_0^t e^{(\alpha-1)(B_s - A_s)} d(B_s - A_s)dt\right] \\
&= E[T] + (\alpha - 1)E\left[\int_0^T \left(\int_s^T dt\right) e^{(\alpha-1)(B_s - A_s)} d(B_s - A_s)\right] \\
&= E[T] + (\alpha - 1)E\left[\int_0^T (T - s)e^{(\alpha-1)(B_s - A_s)} d(B_s) - \int_0^T (T - s)e^{(\alpha-1)(B_s - A_s)} d(A_s)\right] \\
&= E[T] + (\alpha - 1)E\left[\int_0^T (T - s)e^{(\alpha-1)(B_s - A_s)} d(N_s^S) - \int_0^T (T - s)e^{(\alpha-1)(B_s - A_s)} d(N_s^T)\right] \\
&= E[T] + (\alpha - 1)E[(T - S)e^{(\alpha-1)(B_S - A_S)} 1_{\{S \le T\}}] \\
&= E[T] + (\alpha - 1)E[|T - S|e^{(\alpha-1)|(B_{T\wedge S} - A_{T\wedge S})|} 1_{\{S \le T\}}].
\end{aligned}$$

Using the same arguments, we also have

$$CRIG_{\alpha}(S,T) = E\left[\int_0^T e^{(\alpha-1)(A_t-B_t)}dt\right]$$
$$= E[S] + (\alpha-1)E[|T-S|e^{(\alpha-1)|(A_{T\wedge S}-B_{T\wedge S})|}1_{\{T\leq S\}}].$$

Therefore,

$$CRCRIG_{\alpha}(T,S) = E[T] + (\alpha-1)E[|T-S|e^{(\alpha-1)(B_{T\wedge S}-A_{T\wedge S})}1_{\{S\leq T\}}]$$
$$+E[S] + (\alpha-1)E[|T-S|e^{(\alpha-1)(A_{T\wedge S}-B_{T\wedge S})}1_{\{T\leq S\}}]$$
$$= E[T] + E[S] + E[(\alpha-1)|T-S|e^{(\alpha-1)|(B_{T\wedge S}-A_{T\wedge S})|}],$$

as required □

3.2 Conditional Relative Cumulative Residual Information Generating Function between Univariate Nonexplosive Point Processes

A point process over $\Re^+$ can be described by an increasing sequence of random variables or by means of its corresponding counting process.

Definition 3.2.1 An univariate point process is an increasing sequence $T = (T_n)_{n\geq 0}$, $T_0 = 0$ of positive extended random variables, $0 \leq T_1 \leq T_2 \leq ...$, defined on a complete probability space $(\Omega, \Im, P)$. The inequalities are strict unless $T_n = \infty$. If $T_\infty = \lim_{n\to\infty} T_n = \infty$, the point process is called nonexplosive.

Another equivalent way to describe a point process is by a counting process $N = (N_t)_{t\geq 0}$ with

$$N_t^T(w) = \sum_{k\geq 1} 1_{\{T_k(w)\leq t\}},$$

which is, for each realization w, a right continuous step function with $N_0(w) = 0$. As $(N_t)_{t\geq 0}$ and $(T_n)_{n\geq 0}$ carry the same information, the associated counting process is also called point process.

The mathematical description of our observations, at the complete information level, is given by the internal family of sub σ-algebras of $\Im$, denoted by $(\Im_t^T)_{t\geq 0}$, where

$$\Im_t^T = \sigma\{1_{\{T_i>s\}}, i \geq 1, 0 < s < t\}$$

satisfies the Dellacherie conditions of right continuity and completeness.

The point process $(N_t^{\mathbf{T}})_{t\geq 0}$ is adapted to $(\Im_t^{\mathbf{T}})_{t\geq 0}$ and $E[N_t^{\mathbf{T}}|\Im_s^{\mathbf{T}}] \geq N_s^{\mathbf{T}}$ for $s < t$, that is, $N_t^{\mathbf{T}}$ is an uniformly integrable $\Im_t^{\mathbf{T}}$-submartingale. Then, from Doob-Meyer decomposition, there exists a unique right-continuous nondecreasing $\Im_t^{\mathbf{T}}$-predictable and integrable process $(A_t^{\mathbf{T}})_{t\geq 0}$, with $A_0^{\mathbf{T}} = 0$ such that $(M_t^{\mathbf{T}})_{t\geq 0}$, with $N_t^{\mathbf{T}} = A_t^{\mathbf{T}} + M_t^{\mathbf{T}}$, is an uniformly integrable $\Im_t^{\mathbf{T}}$-martingale.

The compensator process is expressed in terms of conditional probabilities, given the available information, and it generalizes the classical notion of hazard. Intuitively, it corresponds to producing whether the failure is going to occur now or not, on the basis of all observations available up to, but not including the present time. Furthermore, the $\Im_t^T$-compensator can be written in the regenerative form as

$$A_t^{\mathbf{T}} = \sum_n 1_{\{T_n\leq t<T_{n+1}\}} A^{(n)}_{(t|T_1,..,T_n)},$$

where each $A^{(n)}_{(t|t_1,..,t_n)}$ is a deterministic function of its arguments. Thus, the compensators are piecewise deterministic functions; each occurrence of a point causes a new function to be selected.

Our aim now is to define a conditional cumulative residual inaccuracy measure between two independent non-explosive point processes, $\mathbf{T}$ and $\mathbf{S}$, and to proceed, as in Bueno and Balakrishnan (2024), to use a superposition process $\mathbf{V} = (V_n)_{n\geq 1}$ with compensator $(A + B)^V_t = A^{\mathbf{T}}_t + B^{\mathbf{S}}_t$, as in Definition 2.2.

At this point, the original compensator process $A^{\mathbf{T}}_t$ is replaced by $(1-\alpha)A^{\mathbf{T}}_t$ with

$$(1-\alpha)A^{\mathbf{T}}_t = \sum_n 1_{\{T_n \leq t < T_{n+1}\}}(1-\alpha)A^{(n)}_{(t|T_1,..,T_n)};$$

$B^{\mathbf{S}}_t$ is replaced by $(1-\alpha)B^{\mathbf{S}}_t$ with

$$1-\alpha)B^{\mathbf{S}}_t = \sum_n 1_{\{S_n \leq t < S_{n+1}\}}(1-\alpha)B^{(n)}_{(t|S_1,..,S_n)};$$

the superposition process $\mathbf{V} = (V_n)_{n\geq 1}$, with compensator $A_t + B_t$, is replaced by $(\alpha - 1)(A_t + B_t)$ with

$$(1-\alpha)(A+B)^{\mathbf{V}}_t = \sum_n 1_{\{V_n \leq t < V_{n+1}\}}(1-\alpha)(A+B)^{(n)}_{(t|V_1,..,S_V)},$$

Following (Bremaud 1981), the original measure P is then replaced by a new measure Q^V, such that Q^V is absolutely continuous with respect to P and

$$\frac{dQ^V}{dP}|_{\Im_t} = L^V_t = \alpha^{\Sigma_{n\geq 1} 1_{\{V_n \leq t\}}} e^{\int_0^t \alpha d(A+B)_s}$$

which, in case where $0 < \alpha < 1$, is an uniformly integrable and locally square integrable martingale under P.

We define a conditional cumulative inaccuracy measure between two univariate non-explosive point processes at any $\Im_t$-stopping time τ, in particular, at time t as follows

Definition 3.2.3 Let $\mathbf{T} = (T_n)_{n\geq 0}$ and $\mathbf{S} = (S_n)_{n\geq 0}$ be point processes with $\Im_t$-compensator processes $(A_t)_{t\geq 0}$ and $(B_t)_{t\geq 0}$, respectively, defined on a complete probability space $(\Omega, \Im, P)$. Let $(V_n)_{n\geq 0}$ be their superposition process. Then, the relative conditional cumulative residual information generating function by time t is

$$CRCRIG_\alpha(N^T_t, N^S_t) = E_{Q^V}\left[\int_0^t e^{(\alpha-1)(B_t - A_t)} dt\right] + E_{Q^V}\left[\int_0^t e^{(\alpha-1)(A_t - B_t)} dt\right]$$

provided the integral exists.

Therefore, we can extend it as

$$CRCRIG_\alpha(N^T, N^S) = E_{Q^V}\left[\int_0^\infty e^{(\alpha-1)(B_t - A_t)} dt\right] + E_{Q^V}\left[\int_0^\infty e^{(\alpha-1)(A_t - B_t)} dt\right].$$

provided the integral exists.

The interpretation of Definition 3.2.3 is given in the following theorem which is proved in Appendix.

Theorem 3.2.4 *Let* $\mathbf{T} = (T_n)_{n\geq 0}$ *and* $\mathbf{S} = (S_n)_{n\geq 0}$ *be nonexplosive point processes and* $\mathbf{V} = (V_n)_{n\geq 0}$ *be their superposition process. Then,*

$$CRCRIG_\alpha(N_t^T, N_t^S) = E_{Q^V}\Bigg\{\sum_{j=1}^{N_t^T}\sum_{k=1}^{N_t^S}\Big[(S_k - T_{j-1}) + (T_j - S_{k-1}) + |T_{j-1} - S_k|(\alpha-1)[e^{-(\alpha-1)|A^j_{T_j\wedge S_k} - B^k_{T_j\wedge S_k}|} - e^{(\alpha-1)|B^k_{T_j\wedge S_k} - A^j_{T_j\wedge S_k}|}]1_{\{S_k\leq T_{j-1}\}} + |S_{k-1} - T_j|(\alpha-1)[e^{-(\alpha-1)|B^k_{T_j\wedge S_k} - A^j_{T_j\wedge S_k}|} - e^{(\alpha-1)|A^j_{T_j\wedge S_k} - B^k_{T_j\wedge S_k}|}]1_{\{T_j\leq S_{k-1}\}}\Big]\Bigg\}$$

and

$$CRCRIG_\alpha(N^T, N^S) = E_{Q^V}\left\{\sum_{n=1}^{\infty}\left[(V_n - V_{n-1}) + (\alpha-1)|V_n - V_{n-1}|(e^{-(\alpha-1)|A^n_{V_n} - B^n_{V_n}|} - (e^{(\alpha-1)|A^n_{V_n} - B^n_{V_n}|})\right]\right\}.$$

4 Preliminaires, Applications and Results

4.1 Preliminaires

Now, we consider another point process $\mathbf{S} = (S_n)_{n\geq 0}$ with $\Im_t$-compensator process $(B_t)_{t\geq 0}$, related to the counting process $N_t^{\mathbf{S}}$ observed at

$$\Im_t^S = \sigma\{1_{\{S_i > s\}}, i \geq 1, 0 < s < t\},$$

with

$$B_t^{\mathbf{S}} = \sum_n 1_{\{S_n < t \leq S_{n+1}\}} B^{(n)}_{(t|S_1,..,S_n)}.$$

Then the following result is adapted from Kwiecinski and Szekli (1991).

Theorem 4.1.1 *If for point processes* $N_t^{\mathbf{T}}$ *and* $N_t^{\mathbf{S}}$*, possibly defined on different probability spaces,*

$$A^{(n)}_{(t_n|t_1,..,t_{n-1})} \leq B^{(n)}_{(t_n|s_1,..,s_{n-1})}, \quad n \geq 1,$$

for $\mathbf{s} \leq \mathbf{t}$*, coordenatewise and for* $P^{\mathbf{T}}$ *almost sure, then there exists a probability space* $(\widetilde{\Omega}, \widetilde{\Im}, \widetilde{P})$ *and point processes* $N_t^{\widetilde{\mathbf{T}}}$ *and* $N_t^{\widetilde{\mathbf{S}}}$ *on it such that* $N_t^{\widetilde{\mathbf{T}}} =^{st} N_t^{\mathbf{T}}$, $N_t^{\widetilde{\mathbf{S}}} =^{st} N_t^{\mathbf{T}}$ *and* $\widetilde{P}(N_t^{\widetilde{\mathbf{T}}} \leq N_t^{\widetilde{\mathbf{S}}},\ \ t \geq 0) = 1$.

At this point, we consider an additional extension of Kullback-Leibler (KL) divergence, at time t, between $\mathbf{T}$ *and* $\mathbf{S}$ *defined possibly on different probability spaces.*

Definition 4.1.2 Let $\mathbf{T} = (T_n)_{n\geq 0}$ be a point process defined on a complete probability space $(\Omega, \Im, Q_1)$, with $\Im_t$-compensator process $(A_t)_{t\geq 0}$, and $\mathbf{S} = (S_n)_{n\geq 0}$ be another point process defined on a complete probability space $(\Omega, \Im, Q_2)$, with $\Im_t$-compensator process $(B_t)_{t\geq 0}$. Then, the KL divergence, at time t, between $\mathbf{T}$ and $\mathbf{S}$ is

$$J_t^C(\mathbf{T}, \mathbf{S}) = E_{Q_1}\left[\int_0^t A_s ds\right] + E_{Q_2}\left[\int_0^t B_s ds\right].$$

4.2 Applications and Results

In what follows we provide several applications of the developed theory, including general repair processe, minimal repair point processes, coherent systems, Markov-modulated Poisson processes and Markov chains. In all cases we present the inequality

$$J_t^C(\mathbf{T}, \mathbf{S}) \leq J_t^C(\mathbf{T}, \mathbf{R}).$$

As $J_t^C(\mathbf{T}, \mathbf{S})$ can be view as a metric between the process $\mathbf{T}$ and the process $\mathbf{S}$, the inequality indicates that the distance between $\mathbf{T}$ and $\mathbf{S}$ is lower than the distance between $\mathbf{T}$ and $\mathbf{R}$ providing a better approximation for $\mathbf{T}$. These include to choose a better repair police, a small parameter distribution, a small set of infinitesimal characteristics of a Markov chain, ...

4.2.1 Application to a General Repair Process

Let $\mathbf{T} = (T_n)_{n\geq 1}$, with $T_0 = 0$, be a point process defined in a complete probability space $(\Omega, \Im, P)$, observed with the history

$$\Im_t^{\mathbf{T}} = \sigma\{1_{\{T_n>s\}}, 0 \leq s < t, n = 1, 2, ...\}$$

satisfying Dellacherie's conditions of right continuity and completeness, describing failures times of a system, with lifetime T, at which instantaneous repairs are carried out. We assume that $T_n, n \geq 1$, are totally inaccessible $\Im_t^{\mathbf{T}}$-stopping times.

Let $(N_t^{\mathbf{T}})_{t\geq 0}$ be the corresponding point process defined as

$$N_t^{\mathbf{T}} = \Sigma_{n=1}^{\infty} 1_{\{T_n \leq t\}}$$

and $(A_t^{\mathbf{T}})_{t\geq 0}$ be the $\Im_t^{\mathbf{T}}$-compensator process of $(N_t)_{t\geq 0}$ in the regenerative form as

$$A_t^{\mathbf{T}} = \sum_n 1_{\{T_n<t\leq T_{n+1}\}} A^{(n)}_{(t|T_1,..,T_n)}.$$

To model varying degrees of repairs, we assume that the nth repair has the effect that the distribution of failures is that of an unfailed item of age $D_n > 0$, where $(D_n)_{n\geq 0}$, with $D_0 = 0$, is a sequence of nonnegative random variables such that D_n is $\Im_{T_n}^{\mathbf{T}}$-measurable. Therefore, under repairs, the general term of the new $\Im_t^{\mathbf{T}}$-compensator of $(N_t)_{t\geq 0}$ is $D_t^{(n)} = A^{(n)}_{t-T_n+D_n}$. In the case when $D_n = T_n$ for all $n \in \mathbf{N}$, $(N_t)_{t\geq 0}$ is a nonhomogeneous Poisson process with general term compensator function $D_t^{(n)} = A_t^{(n)}$. When $D_n = 0$ for all $n \in \mathbf{N}$, $(N_t)_{t\geq 0}$ is a renewal process with general term compensator function $D_t^{(n)} = A^{(n)}_{t-T_n}$.

We now proceed with a change of probability measure, say, from P to Q, following Girsanov theorem, see (Bremaud 1981), in which the occurrence times are retained. We can write

$$D_t^{(n)} = A^{(n)}_{t-T_n+D_n} = \int_{(T_n-D_n)\vee 0}^{t} dA_s,$$

and then establish the theorem.

Theorem 4.2.1.1 *The process $(L_t)_{t\geq 0}$ defined by*

$$L_t = (1_{\{T>(T_n-D_n)\vee 0\}})^{1_{\{T\leq t\}}} e^{\int_0^t dA_s - \int_{(T_n-D_n)\vee 0}^t dA_s} = (1_{\{T>(T_n-D_n)\vee 0\}})^{1_{\{T\leq t\}}} e^{A^{(n)}_{(T_n-D_n)\vee 0}}$$

is an integrable $\Im_t^{\mathbf{T}}$-martingale, with $E[L_t] = 1$.

Proof We have to prove that L_t is $\Im_t^{\mathbf{T}}$ adapted, the martingale property $E[L_t|\Im_s^{\mathbf{T}}] = L_s$ and L_t is integrable.

Clearly, as T is $\Im_t$- measurable and A_t is $\Im_t$- predictable we have that L_t is $\Im_t^{\mathbf{T}}$ adapted.

We have to prove that, for all $T_n \leq s < t < T_{n+1}$, $E[L_t|\Im_s^{\mathbf{T}}] = L_s$, which is equivalent to proving that

$$E[L_t - L_s|\Im_s^{\mathbf{T}}] = E\left\{\left[(1_{\{T>(T_n-D_n)\vee 0\}})^{1_{\{T\leq t\}}} - (1_{\{T>(T_n-D_n)\vee 0\}})^{1_{\{T\leq s\}}}\right] e^{A^{(n)}_{(T_n-D_n)\vee 0}}|\Im_s^{\mathbf{T}}\right\} = 0.$$

Observe that if $T \leq s$, then $T \leq t$ and $E[L_t - L_s|\Im_s^{\mathbf{T}}] = 0$. Also, if $T > t$, then $T > s$ and $E[L_t - L_s|\Im_s^{\mathbf{T}}] = 0$.

If $T_n < s < T \leq t \leq T_{n+1}$, we observe that

$$\{N_{t\wedge s} = k\} \cap \{T_n \leq s < T_{n+1}\} = \{N_{t\wedge T_n} = k\} \cap \{T_n \leq s < T_{n+1}\}.$$

Therefore, for any generator U of $\Im_s$, there exists a generator V of $\Im_{T_n}$ such that

$$U \cap \{T_n \leq s < T_{n+1}\} = V \cap \{T_n \leq s < T_{n+1}\}$$

and, as in Theorem 32, Appendix 2, of Bremaud (1981), the information before s is the same as the information before T_n. Also, observe that if, $T > T_n$, $(1_{\{T>(T_n-D_n)\vee 0\}}) = 1$. Hence,

$$E\left[L_t|\Im_s^{\mathbf{T}}\right] = E\left[(1_{\{T>(T_n-D_n)\vee 0\}})e^{A^{(n)}_{(T_n-D_n)\vee 0}}|\Im_s^{\mathbf{T}}\right] = E\left[e^{A^{(n)}_{(T_n-D_n)\vee 0}}|\Im_s^{\mathbf{T}}\right]$$
$$= E\left[e^{A^{(n)}_{(T_n-D_n)\vee 0}}|\Im_{T_n}^{\mathbf{T}}\right] = e^{A^{(n)}_{(T_n-D_n)\vee 0}} = L_s.$$

The third equality follows because $e^{A^{(n)}_{(T_n-D_n)\vee 0}}$ is $\Im_{T_n}^{\mathbf{T}}$-measurable. Also, as $\Im_0^{\mathbf{T}} = \{\Omega, \emptyset\}$, we have $E[L_t] = E[L_0] = E[e^{A_0}] = 1$ and L_t is integrable. Hence, the theorem. □

Theorem 4.2.1.2 *Under the probability measure Q defined by the Radon-Nikodym derivative*

$$\frac{dQ}{dP}|_{\Im_\infty} = L_\infty = (1_{\{T>T_n-D_n\vee 0\}})e^{A^{(n)}_{(T_n-D_n)\vee 0}},$$

$(N_t^{(n)} - A^{(n)}_{t-T_n+D_n})_{t\geq 0}$ *is an* $\Im_t^{\mathbf{T}}$*-martingale, where* $N_t^{(n)}$ *is* N_t *restricted to* $[T_n, T_{n+1})$.

Proof Clearly, we have

$$E_Q[A(t - T_n + D_n)] = E_Q\left[\int_{(T_n-D_n)\vee 0}^{t} dA_s\right] = E\left[(1_{\{T>(T_n-D_n)\vee 0\}})e^{A^{(n)}_{(T_n-D_n)\vee 0}}\int_{T_n-D_n}^{t} dA_s\right]$$
$$= E\left[\int_{(T_n-D_n)\vee 0}^{t}(1_{\{T>(T_n-D_n)\vee 0\}})e^{A^{(n)}_{(T_n-D_n)\vee 0}}dA_s\right]$$
$$= E\left[\int_{(T_n-D_n)\vee 0}^{t}(1_{\{T>(T_n-D_n)\vee 0\}})e^{A^{(n)}_{(T_n-D_n)\vee 0}}dN_s\right]$$

$$= E[(1_{\{T>(T_n-D_n)\vee 0\}})e^{A^{(n)}_{(T_n-D_n)\vee 0}}1_{\{(T_n-D_n)\vee 0<T\leq t\}}] = E_Q[1_{\{T\leq t\}}],$$

and conclude that, under Q, $(N_t^{(n)} - A^{(n)}_{t-T_n+D_n})_{t\geq 0}$ is an $\Im_t^{\mathbf{T}}$-martingale, where $N_t^{(n)}$ is N_t restrict to $[T_n, T_{n+1})$.

Thus, the proof gets completed and, under Q, $A^{(n)}_{t-T_n+D_n}$ is the unique general term of the $\Im^{\mathbf{T}}_t$-compensator of $(N_t)_{t\geq 0}$.

To proceed with a repair policy, we must look for the most convenient type of repair we should use to best approximate the ideal system performance.

Let us now consider the point process $\mathbf{S} = (S_n)_{n\geq 1}$, with $S_0 = 0$, corresponding to the counting process $(N_t)_{t\geq 0}$, resulting from repairing the process $\mathbf{T}$, characterized by a sequence of $\Im^{\mathbf{T}}_{T_n}$-measurable nonnegative random variables $(U_n)_{n\geq 1}$, with $U_0 = 0$, and $\Im^{\mathbf{T}}_t$-compensator process

$$B_t = \sum_n 1_{\{T_n < t \leq T_{n+1}\}} B^{(n)}_{(t|T_1,..,T_n)},$$

with $B^{(n)}_{(t|T_1,..,T_n)} = A^{(n)}_{t-T_n+U_n}$, under a probability measure Q_1.

Further, let us consider another process $\mathbf{R} = (R_n)_{n\geq 1}$, $R_0 = 0$, corresponding to the counting process $(N_t)_{t\geq 0}$, resulting from repairing the process $\mathbf{T}$, characterized by a sequence of $\Im^{\mathbf{T}}_{T_n}$-measurable nonnegative random variables $(V_n)_{n\geq 1}$, with $V_0 = 0$, and $\Im^{\mathbf{T}}_t$-compensator process

$$C_t = \sum_n 1_{\{T_n < t \leq T_{n+1}\}} C^{(n)}_{(t|T_1,..,T_n)},$$

with $C^{(n)}_{(t|T_1,..,T_n)} = A^{(n)}_{t-T_n+V_n}$, under a probability measure Q_2.

Clearly, if

$$P(\{w : U_n(w) \leq V_n(w), \ \ n = 1, 2, 3, ...\}) = 1,$$

as $A^{(n)}_t$ is nondecreasing, we will have

$$P(\{w : A^{(n)}_{t-T_n+U_n} \leq A^{(n)}_{t-T_n+V_n}, \ \ n = 1, 2, 3, ...\}) = 1.$$

From Theorem 4.1.1 we have $\widetilde{P}(N^{\widetilde{\mathbf{T}}}_t \leq N^{\widetilde{\mathbf{S}}}_t, \ \ t \geq 0) = 1$, and we then finally obtain

$$J^C_t(\mathbf{T}, \mathbf{S}) \leq J^C_t(\mathbf{T}, \mathbf{R}).$$

To motivate this result, let $U_n = 0$ for all n (a renewal process) and $V_n = T_n$ for all n (a minimal repair process). As $P(\{w : U_n(w) \leq V_n(w), \ \ n = 1, 2, 3, ...\}) = 1$, we can conclude that the renewal process produces a better approximation for $\mathbf{T}$ than a minimal repair process. □

4.2.2 Application to a Minimal Repair Process

A repair is minimal if the intensity λ^T_t is not affected by the occurrence of failures; in other words, we cannot determine the failure time points from the observation of λ^T_t. Formally, we have the following definition.

Definition 4.2.2.1 Let $T = (T_n)_{n\geq 0}$ be a point process with an integrable point process N^T and corresponding $\Im_t$-intensity $(\lambda^T_t)_{t\geq 0}$. Let $\Im^{\lambda^T}_t = \sigma\{\lambda^T_s, 0 \leq s \leq t\}$ be the filtration generated by λ^T. Then, the point process T is said to be a minimal repair process (MRP) if none of the variables T_n, $n \geq 0$, for which $P(T_n < \infty) > 0$, is an $\Im^{\lambda^T}_t$-stopping time.

If T is a non-homogeneous Poisson process, then $\lambda_t = \lambda(t)$ is a time dependent deterministic function, and it means that the age is not changed as a result of a failure. Here, $\Im^{\lambda^T}_t = \{\Omega, \emptyset\}$ for all $t \in \Re^+$, and the failure times T_n are not $\Im^{\lambda^T}_t$-stopping times.

Let $(T_n)_{n\geq 0}$ be a Weibull process with parameters β and θ_1, that is, we consider the ordered lifetimes $T_1, ..., T_n$ with a conditional reliability function given by

$$\overline{G_i}(t_i|t_1, ..., t_{i-1}) = \exp\left\{-\left(\frac{t_i}{\theta_1}\right)^\beta + \left(\frac{t_{i-1}}{\theta_1}\right)^\beta\right\}$$

for $0 \leq t_{i-1} < t_i$, where t_i are the ordered observations.

The $\Im^T$-compensator process is then

$$A^{(n)}_{(t|t_1,..,t_n)} = \sum_{j=1}^{n}\left[\left(\frac{t_j}{\theta_1}\right)^\beta - \left(\frac{t_{j-1}}{\theta_1}\right)^\beta\right] + \left[\left(\frac{t}{\theta_1}\right)^\beta - \left(\frac{t_n}{\theta_1}\right)^\beta\right] = \left(\frac{t}{\theta_1}\right)^\beta.$$

Now, suppose we observe the Weibull process $(S_n)_{n\geq 1}$, with parameters β and θ_2, through the family $\mathbf{G}^{\mathbf{S}} = (G^{\mathbf{S}}_t)_{t\geq 0}$, where

$$G^{\mathbf{S}}_t = \sigma\{1_{\{S_n>s\}},\ \ 0 \leq s < t\}$$

satisfies Dellacherie's conditions of right continuity and completeness asserted by the experimenter for $(T_n)_{n\geq 0}$. We can then calculate $J^C_t(S, T)$, where $(N^S_t - B_t)_{t\geq 0}$ is an uniformly integrable martingale, as

$$J^C_t(\mathbf{S}, \mathbf{T}) = E\left[\int_0^t A_s ds + \int_0^t B_s ds\right] = E\left[\int_0^t \left(\frac{s}{\theta_1}\right)^\beta ds + \int_0^t \left(\frac{s}{\theta_2}\right)^\beta ds\right]$$

$$= \frac{t^{\beta+1}}{\beta+1}\left(\frac{\theta_1^\beta + \theta_2^\beta}{\theta_1^\beta\theta_2^\beta}\right).$$

Also, let $(R_n)_{n\geq 1}$ be a Weibull process, with parameters β and θ_3, observed through the family $\mathbf{G}^{\mathbf{R}} = (G^{\mathbf{R}}_t)_{t\geq 0}$, where

$$G^{\mathbf{R}}_t = \sigma\{1_{\{R_n>s\}},\ \ 0 \leq s < t\}$$

satisfies Dellacherie's conditions of right continuity and completeness asserted by the experimenter for $(T_n)_{n\geq 0}$. We then have

$$J^C_t(\mathbf{R}, \mathbf{T}) = \frac{t^{\beta+1}}{\beta+1}\left(\frac{\theta_1^\beta + \theta_3^\beta}{\theta_1^\beta\theta_3^\beta}\right).$$

The $\mathbf{G}^{\mathbf{R}}$-compensator process of $(R_n)_{n\geq 0}$ is $C^{(n)}_{(t|t_1,..,t_n)} = (\frac{t}{\theta_3})^\beta$.

If we assume $\theta_3 \leq \theta_2$, we then have $(\frac{t}{\theta_2})^\beta \leq (\frac{t}{\theta_3})^\beta$ and so, in this case, Theorem 4.1.1 holds, that is,

$$J^C_t(\mathbf{S}, \mathbf{T}) \leq J^C_t(\mathbf{R}, \mathbf{T}).$$

4.2.3 Application to a Parallel System Minimally Repaired at Component Level

Let $(U_n)_{n\geq 1}$ be a sequence of random variables independent and identically distributed as exponential, with scale parameter α, through the family $\mathbf{F} = (\Im_t)_{t\geq 0}$, where

$$\Im_t = \sigma\{1_{\{U_n>s\}},\ \ n \geq 1, 0 \leq s < t\}$$

satisfies Dellacherie conditions of right continuity and completeness.

Let $(T_n)_{n\geq 1}$ be the point process and the corresponding counting process

$$N_t = \Sigma_{n=1}^{\infty} 1_{\{T_n \leq t\}}$$

defined by $T = T_1 = U_1 \vee U_2 = \max\{U_1, U_2\}$ and $T_{n+1} = T_n + U_{n+2}, n \geq 1$.

The intensity of N_t is $\lambda_t = \alpha 1_{\{U_1 \wedge U_2 \leq t\}}$ and so $\Im_t^{\lambda} = \sigma\{1_{\{U_1 \wedge U_2 > t\}},\ \ 0 \leq s < t\}$. Clearly, T_1 is not $\Im_t^{\lambda}$- measurable and $T_n = T_1 + U_3 + ... + U_{n+1}, n \geq 2$, are not $\Im_t^{\lambda}$- measurable. Therefore, N_t is a minimal repair point process.

The conditional survival function of T_1 is

$$P(T_1 > t|\Im_t) = e^{-\alpha(t-U_1 \wedge U_2)^+},$$

where $U_1 \bigwedge U_2 = \min\{U_1, U_2\}$. However, its physical lifetime is identically distributed as

$$U_2 \vee U_1 - U_2 \wedge U_1 = |U_2 - U_1|,$$

where $U_1 \bigvee U_2 = \max\{U_1, U_2\}$, which has an exponential distribution with parameter α. Furthermore, the interarrival times $T_{n+1} - T_n = U_{n+2}, n \geq 1$, are independent and identically distributed as exponential with parameter α.

Therefore, the general term of the compensator function is

$$A^{(n)}_{(t|t_1,..,t_n)} = \alpha \times t.$$

Next, suppose we observe a sequence of random variables $(V_n)_{n\geq 1}$ of independent and identical exponential random variables with parameter β through the family $\mathbf{G}^{\mathbf{X}} = (G_t^{\mathbf{X}})_{t\geq 0}$, where

$$G_t^{\mathbf{V}} = \sigma\{1_{\{V_n > s\}},\ \ n \geq 1,\ \ 0 \leq s < t\}$$

satisfies Dellacherie conditions of right continuity and completeness to generate minimally repaired two-component parallel systems, as above, with lifetimes $\mathbf{S}$. We can then calculate $J_t^C(\mathbf{S}, \mathbf{T})$, where $(N_t^{\mathbf{S}} - B_t)_{t\geq 0}$ is an uniformly integrable martingale and the general term of the compensator function is

$$B^{(n)}_{(s|s_1,..,s_{n-1})} = \beta \times s.$$

Alternatively, we can observe a sequence of random variables $(W_n)_{n\geq 1}$ of independent and identically exponential random variables with parameter γ through the family $\mathbf{G}^{\mathbf{W}} = (G_t^{\mathbf{W}})_{t\geq 0}$, where

$$G_t^{\mathbf{W}} = \sigma\{1_{\{W_n > s\}},\ \ n \geq 1,\ \ 0 \leq s < t\}$$

satisfies Dellacherie conditions of right continuity and completeness to generate minimally repaired two-component parallel systems, as above, with lifetimes $\mathbf{R}$. We can then calculate $J_t^C(\mathbf{R}, \mathbf{T})$, where $(N_t^{\mathbf{R}} - C_t)_{t\geq 0}$ is an uniformly integrable martingale and the general term of the compensator function is

$$C^{(n)}_{(s|s_1,..,s_{n-1})} = \gamma \times s.$$

Here, if $\beta < \gamma$, we have

$$B^{(n)}_{(s|s_1,..,s_{n-1})} = \beta \times s \leq \gamma \times s = C^{(n)}_{(s|s_1,..,s_{n-1})}$$

and we can then use Theorem 4.1.1 to conclude that

$$J_t^C(\mathbf{S}, \mathbf{T}) \leq J_t^C(\mathbf{R}, \mathbf{T}).$$

4.2.4 Application to a Markov-Modulated Poisson Process

A Poisson process can be generalized by replacing the constant intensity with a randomly varying intensity, which takes one of n values $\lambda_i,\ 0<\lambda_i<\infty,\ i\in S=\{1,2,...,n\}, n\in\mathbf{N}$. The changes are driven by a homogeneous Markov chain $X=(X_t)_{t\geq 0}$ with values in S, with infinitesimal parameters q_i, as the rate to leave state i, and q_{ij} as the rate to reach state j from state i.

The point process $\mathbf{T}=(T_n)_{n\geq 1}$ corresponding to the counting process $N=(N_t)_{t\geq o}$, with

$$N_t=\sum_{n=1}^{\infty}1_{\{T_n\leq t\}},$$

has a stochastic intensity λ_{X_t} with respect to the filtration $(\Im_t)_{t\geq 0}$, generated by N and X:

$$\sigma\{N_s,X_s,\ \ 0\leq s\leq t\}.$$

Then, N is said to be a Markov-modulated Poisson process with smooth semi-martingale representation

$$N_t=\int_0^t\lambda_{X_s}ds+M_t,$$

where $(M_t)_{t\geq 0}$ is a $\Im_t$-martingale and $\lambda_{X_s}=\sum_{j=1}^{n}\lambda_j 1_{\{X_s=j\}}$.

Furthermore, the indicator of state j at time t, $1_{\{X_t=j\}}$, also has its smooth semi-martingale representation as

$$1_{\{X_t=j\}}=1_{\{X_0=j\}}+\int_0^t\sum_{i=1}^{n}1_{\{X_t=i\}}q_{ij}ds+M_t(j),$$

where $(M_t(j))_{t\geq 0}$ is a zero mean $\Im_t$-martingale.

Hence,

$$N_t=\int_0^t\sum_{j=1}^{n}\sum_{i=1}^{n}\lambda_j 1_{\{X_t=i\}}q_{ij}ds+M_t.$$

The general term of the compensator function is

$$A^{(n)}_{(t|t_1,..,t_{n-1})}=\int_0^t\sum_{j=1}^{n}\sum_{i=1}^{n}\lambda_j 1_{\{X_t=i\}}q_{ij}ds.$$

The cumulative residual inaccuracy measure, at time t, $J_t^C(\mathbf{T},\mathbf{T}^*)$, between the counting process $(N_t^*)_{t\geq 0}$ related a Markov chain $(X_t^*)_{t\geq 0}$ with infinitesimal matrix $Q^*=[q_{ij}^*]$, $i,j\in\mathbf{N}_+$, asserted by the experimenter's information of the true Markov chain $(X_t)_{t\geq 0}$ having infinitesimal parameter $Q=[q_{ij}]$, $i,j\in\mathbf{N}_+$ is given by

$$J_t^C(\mathbf{T},\mathbf{T}^*)=E\left[\int_0^t\sum_{j=1}^{n}\sum_{i=1}^{n}\lambda_j 1_{\{X_t=i\}}q_{ij}ds\right]+E\left[\int_0^t\sum_{j=1}^{n}\sum_{i=1}^{n}\lambda_j 1_{\{X_t=i\}}q^*_{ij}ds\right].$$

The general term of the compensator function, where $(N_t^*-B_t)$ is a 0 mean $\Im_t$-martingale, is

$$B^{(n)}_{(t|t_1,..,t_{n-1})}=\int_0^t\sum_{j=1}^{n}\sum_{i=1}^{n}\lambda_j 1_{\{X_t=i\}}q^*_{ij}ds.$$

Let us now consider another process $(N_t^{**})_{t\geq 0}$ related to a Markov chain $(X_t^{**})_{t\geq 0}$ with infinitesimal matrix $Q^{**} = [q_{ij}^{**}]$, $i, j \in \mathbf{N}_+$, asserted by the experimenter's information of the true Markov chain $(X_t)_{t\geq 0}$. Now, the general term of the compensator function, where $(N_t^{**} - C_t)$ is a 0 mean $\Im_t$-martingale, is

$$C^{(n)}_{(t|t_1,..,t_{n-1})} = \int_0^t \sum_{j=1}^n \sum_{i=1}^n \lambda_j 1_{\{X_t=i\}} q_{ij}^{**} ds.$$

If we assume that $\max_{i,j\in\mathbf{N}_+}\{q_{ij}^*\} \leq \min_{i,j\in\mathbf{N}_+}\{q_{ij}^{**}\}$, then

$$B^{(n)}_{(t_n|t_1,..,t_{n-1})} \leq C^{(n)}_{(t_n|s_1,..,s_{n-1})}, \quad n \geq 1,$$

and we can then use Theorem 4.1.1 to conclude that

$$J_t^C(\mathbf{T}, \mathbf{T}^*) \leq J_t^C(\mathbf{T}, \mathbf{T}^{**}).$$

4.2.5 Application to a Markov Chain

Let $(X_t)_{t\geq 0}$ be a Markov chain defined on a probability space $(\Omega, \Im, P)$ and adapted to some history $(\Im_t)_{t\geq 0}$. The observations are through its internal history

$$\Im_t^X = \sigma\{X_s, s \leq t\}$$

for all $t \geq 0$, and $\Im_t^X \subseteq \Im_t$ for all $t \geq 0$. Then, $\Im_\infty^X$ records all the events linked to the process $(X_t)_{t\geq 0}$.

The $\Im_t$-Markov chain is associated with a sequence of its sojourn times $(T_{n+1} - T_n)_{n\geq 0}$, with $T_0 = 0$, and its infinitesimal characteristics $Q = [q_{i,j}]$, $(i, j) \in \mathbf{N}^+ X \mathbf{N}^+$. If, for each natural number i, we have

$$q_i = \sum_{j\neq i} q_{i,j} < \infty,$$

the chain is said to be stable and conservative. We set $q_{i,i} = -q_i$.

We are now interested in the cumulative inaccuracy process between point processes $N_t^X(l)$: the number of transitions into state l during the interval $(0, t]$, related to Markov chain's occurrence times.

The $\Im_t$-compensator of $N_t(l)$ is

$$\int_0^t \sum_{i\neq l} q_{i,l} 1_{\{X_u=i\}} du,$$

provided

$$\sum_{i\neq l} \int_0^t q_{i,l} 1_{\{X_u=i\}} du < \infty.$$

The general term of the compensator function is

$$A^{(n)}_{(t|t_1,..,t_{n-1})} = \int_0^t \sum_{i\neq l}^n q_{i,l} 1_{\{X_u=i\}} du.$$

To evaluate the cumulative inaccuracy measure, at time t, between the counting process $N_t^*(l)$, related to Markov chain $(X_t^*)_{t\geq 0}$ with infinitesimal matrix $Q^* = [q_{i,j}^*]$, $i, j \in \mathbf{N}_+$, asserted by the experimenter's information of the true Markov chain $(X_t)_{t\geq 0}$ that has infinitesimal matrix $Q = [q_{i,j}]$, $i, j \in \mathbf{N}_+$, with counting process $N_t(l)$, using Definition 3.1.5, we have

$$J_t^C(\mathbf{X}, \mathbf{X}^*) = E\left[\int_0^t \left(\int_0^s \sum_{i\neq l} q_{i,l} 1_{\{X_u=i\}} du\right) ds\right] + E\left[\int_0^t \left(\int_0^s \sum_{i\neq l} q^*_{i,l} 1_{\{X_u=i\}} du\right) ds\right].$$

Now, the general term of the compensator function $\varphi_B(\mathbf{t})$, where $(N_t^* - B_t)$ is a 0 mean $\Im_t$-martingale, is

$$B^{(n)}_{(t|t_1,..,t_{n-1})} = \int_0^t \sum_{i\neq l}^n q^*_{i,l} 1_{\{X_u=i\}} du.$$

Let us now consider another process $(N_t^{**})_{t\geq 0}$ related to a Markov chain $(X_t^{**})_{t\geq 0}$ with infinitesimal matrix $Q^{**} = [q^{**}_{ij}]$, $i, j \in \mathbf{N}_+$, asserted by the experimenter's information of the true Markov chain $(X_t)_{t\geq 0}$. Now, the general term of the compensator function, where $(N_t^{**} - C_t)$ is a 0 mean $\Im_t$-martingale, is

$$C^{(n)}_{(t|t_1,..,t_{n-1})} = \int_0^t \sum_{i\neq l}^n q^{**}_{i,l} 1_{\{X_u=i\}} du.$$

If we assume that $\max_{i\in\mathbf{N}_+}\{q^*_{il}\} \leq \min_{i\in\mathbf{N}_+}\{q^{**}_{il}\}$, then

$$B^{(n)}_{(t_n|t_1,..,t_{n-1})} \leq C^{(n)}_{(t_n|s_1,..,s_{n-1})}, \quad n \geq 1,$$

and we can then use Theorem 4.1.1 to conclude that

$$J_t^C(\mathbf{X}, \mathbf{X}^*) \leq J_t^C(\mathbf{X}, \mathbf{X}^{**}).$$

5 Concluding Remarks

In this work, using a martingale approach, we have extended a relative cumulative residual information generating measure, $(RCRIG_\alpha(T, S))$, to a conditional one $CRCRIG_\alpha(T, S))$. Using the same technique, we get a conditioned Kullback-Leibler $(J^C(S, T))$ divergence measure. Interestingly, this concept is identical to the cumulative residual inaccuracy measure of Bueno and Balakrishnan (2022). The conditional relative cumulative residual information generating measure generates the Kulback measure through its first derivative at $\alpha = 1$.

We have also extended the conditional relative cumulative residual information generating measure of univariate non-explosive point process $(CRCRIG_\alpha(N^T, N^S))$ to a dynamic conditional one. Then, we have made of conditioned Kulback-Leibler divergence to compare measures between two non-explosive point processes through stochastic inequalities between compensator processes related to the respective counting processes. Several applications and examples have been given, especially in reliability theory. In our future work, we plan to make use of conditioned classes of distributions, such as multivariate increasing (decreasing) failure rate distributions given the observed past, $\Im_t$, to provide sufficient conditions for such inequalities to hold. We are currently working in this directions and hope to report the corresponding findinges in a future paper.

Appendix

Theorem 3.2.4 *Let* $\mathbf{T} = (T_n)_{n\geq 0}$ *and* $\mathbf{S} = (S_n)_{n\geq 0}$ *be nonexplosive point processes and* $\mathbf{V} = (V_n)_{n\geq 0}$ *be their superposition process. Then,*

$$CRCRIG_\alpha(N_t^T, N_t^S) = E_{Q^V}\Bigg\{\sum_{j=1}^{N_t^T}\sum_{k=1}^{N_t^S}\Big[(S_k - T_{j-1}) + (T_j - S_{k-1}) + |T_{j-1} - S_k|(\alpha-1)[e^{-(\alpha-1)|A^j_{T_j\wedge S_k} - B^k_{T_j\wedge S_k}|} - e^{(\alpha-1)|B^k_{T_j\wedge S_k} - A^j_{T_j\wedge S_k}|}]1_{\{S_k\leq T_{j-1}\}} + |S_{k-1} - T_j|(\alpha-1)[e^{-(\alpha-1)|B^k_{T_j\wedge S_k} - A^j_{T_j\wedge S_k}|} - e^{(\alpha-1)|A^j_{T_j\wedge S_k} - B^k_{T_j\wedge S_k}|}]1_{\{T_j\leq S_{k-1}\}}\Big]\Bigg\}$$

and

$$CRCRIG_\alpha(N^T, N^S) = E_{Q^V}\left\{\sum_{n=1}^{\infty}\Big[(V_n - V_{n-1}) + (\alpha-1)|V_n - V_{n-1}|(e^{-(\alpha-1)|A^n_{V_n} - B^n_{V_n}|} - (e^{(\alpha-1)|A^n_{V_n} - B^n_{V_n}|})\Big]\right\}.$$

Proof First, we observe the partitions of $\mathbb{R}$: $\Omega = \cup_{k=1}^{\infty}(S_{k-1}, S_k]$, $\Omega = \cup_{j=1}^{\infty}(T_{j-1}, T_j]$. Thence $\Omega = \cup_{k=1}^{\infty}\cup_{j=1}^{\infty}(S_{k-1}\vee T_{j-1}, S_k\wedge T_j)$ is a partition of $\mathbb{R}$.

We let $(\tau_n^{\mathbf{T}})_{n\geq 0}$ be an increasing sequence of $\Im_t$-stopping times as the localizing sequence of the stopped martingale $(N^{\mathbf{T})}_{t\wedge\tau_n^{\mathbf{T}}} - A_{t\wedge\tau_n^{\mathbf{T}}})_{t\geq 0}$ and let $(\tau_n^{\mathbf{S}})_{n\geq 0}$ be an increasing sequence of $\Im_t$-stopping times as the localizing sequence of the stopped martingale $(N^{\mathbf{S}}_{t\wedge\tau_n^{\mathbf{S}}} - B_{t\wedge\tau_n^{\mathbf{S}}})_{t\geq 0}$ and then apply the Optimal Sampling Theorem.

Note that $\tau_n = \tau_n^{\mathbf{T}}\vee\tau_n^{\mathbf{S}}$ is also an $\Im_t$-stopping time and that the point process $(S_k)_{k\geq 0}$ define a partition of $\Re^+$, that is, $\Re^+ = \cup_{k=0}^{\infty}(S_{k-1}, S_k]$. Therefore, we can write

$$E_{Q^V}\left[\int_0^t e^{(\alpha-1)(B_t - A_t)}dt\right] = E_{Q^V}\left[\sum_{j=1}^{N^T_{\tau_n}}\sum_{k=1}^{N^S_{\tau_n}}\int_{S_{k-1}\vee T_{j-1}}^{S_k\wedge T_j\wedge\tau_n} e^{(\alpha-1)(B_t^k - A_t^j)}dt\right]$$

$$= E_{Q^V}\left\{\sum_{j=1}^{N^T_{\tau_n}}\sum_{k=1}^{N^S_{\tau_n}}\int_{S_{k-1}\vee T_{j-1}}^{S_k\wedge T_j\wedge\tau_n}[1 + \int_0^s(\alpha-1)e^{(\alpha-1)(B_s^k - A_s^j)}d(B_s^k - A_s^j)]dt\right\}$$

$$= E_{Q^V}\left\{\sum_{j=1}^{N^T_{\tau_n}}\sum_{k=1}^{N^S_{\tau_n}}\left[(S_k\wedge T_j\wedge\tau_n - S_{k-1}\vee T_{j-1}) + \int_{S_{k-1}\vee T_{j-1}}^{S_k\wedge T_j\wedge\tau_n}(\alpha-1)\int_0^s e^{(\alpha-1)(B_s^k - A_s^j)}d(B_s^k - A_s^j)dt\right]\right\}$$

$$= E_{Q^V}\Bigg\{\sum_{j=1}^{N^T_{\tau_n}}\sum_{k=1}^{N^S_{\tau_n}}\Bigg[(S_k\wedge T_j\wedge\tau_n - S_{k-1}\vee T_{j-1}) + \int_0^{S_{k-1}\vee T_{j-1}}\left(\int_{S_{k-1}\vee T_{j-1}}^{S_k\wedge T_j\wedge\tau_n}dt\right)(\alpha-1)e^{(\alpha-1)(B_s^k - A_s^j)}d(B_s^k - A_s^j) + \int_{S_{k-1}\vee T_{j-1}}^{S_k\wedge T_j\wedge\tau_n}\left(\int_s^{S_k\wedge T_j\wedge\tau_n}dt\right)(\alpha-1)e^{(\alpha-1)(B_s^k - A_s^j)}d(B_s^k - A_s^j)\Bigg]\Bigg\}$$

$$= E_{Q^V}\Bigg\{\sum_{j=1}^{N^T_{\tau_n}}\sum_{k=1}^{N^S_{\tau_n}}\Bigg[(S_k \wedge T_j \wedge \tau_n - S_{k-1} \vee T_{j-1}) + \int_0^{S_{k-1}\vee T_{j-1}} (S_k \wedge T_j \wedge \tau_n - S_{k-1} \vee T_{j-1})(\alpha-1)e^{(\alpha-1)(B^k_s - A^j_s)} d(B^k_s - A^j_s) + \int_{S_{k-1}\vee T_{j-1}}^{S_k \wedge T_j} (S_k \wedge T_j \wedge \tau_n - s)(\alpha-1)e^{(\alpha-1)(B^k_s - A^j_s)} d(B^k_s - A^j_s)\Bigg]\Bigg\}$$

$$= E_{Q^V}\Bigg\{\sum_{j=1}^{N^T_{\tau_n}}\sum_{k=1}^{N^S_{\tau_n}}\Bigg[(S_k \wedge T_j \wedge \tau_n - S_{k-1} \vee T_{j-1}) + \int_0^{S_{k-1}\vee T_{j-1}} (S_k \wedge T_j \wedge \tau_n - S_{k-1} \vee T_{j-1})(\alpha-1)e^{(\alpha-1)(B^k_s - A^j_s)} dB^k_s - \int_0^{S_{k-1}\vee T_{j-1}} (S_k \wedge T_j \wedge \tau_n - S_{k-1} \vee T_{j-1})(\alpha-1)e^{(\alpha-1)(B^k_s - A^j_s)} dA^j_s + \int_{S_{k-1}\vee T_{j-1}}^{S_k \wedge T_j \wedge \tau_n} (S_k \wedge T_j \wedge \tau_n - s)(\alpha-1)e^{(\alpha-1)(B^k_s - A^j_s)} dB^k_s - \int_{S_{k-1}\vee T_{j-1}}^{S_k \wedge T_j \wedge \tau_n} (S_k \wedge T_j \wedge \tau_n - s)(\alpha-1)e^{(\alpha-1)(B^k_s - A^j_s)} dA^j_s\Bigg]\Bigg\}$$

$$= E_{Q^V}\Bigg\{\sum_{j=1}^{N^T_{\tau_n}}\sum_{k=1}^{N^S_{\tau_n}}\Bigg[(S_k \wedge T_j \wedge \tau_n - S_{k-1} \vee T_{j-1}) + (S_k \wedge T_j \wedge \tau_n - S_{k-1} \vee T_{j-1})(\alpha-1)e^{(\alpha-1)(B^k_{S_k} - A^j_{S_k})} 1_{\{S_k \le S_{k-1}\vee T_{j-1}\}} - (S_k \wedge T_j \wedge \tau_n - S_{k-1} \vee T_{j-1})(\alpha-1)e^{(\alpha-1)(B^k_{T_j} - A^j_{T_j})} 1_{\{T_j \le S_{k-1}\vee T_{j-1}\}} + (S_k \wedge T_j \wedge \tau_n - S_k)(\alpha-1)e^{(\alpha-1)(B^k_{S_k} - A^j_{S_k})} 1_{\{S_{k-1}\vee T_{j-1} \le S_k \le S_k \wedge T_j \wedge \tau_n\}} - (S_k \wedge T_j \wedge \tau_n - T_j)(\alpha-1)e^{(\alpha-1)(B^k_{T_j} - A^j_{T_j})} 1_{\{S_{k-1}\vee T_{j-1} \le T_j \le S_k \wedge T_j \wedge \tau_n\}}\Bigg]\Bigg\}$$

$$= E_{Q^V}\Bigg\{\sum_{j=1}^{N^T_{\tau_n}}\sum_{k=1}^{N^S_{\tau_n}}\Bigg[(S_k \wedge T_j \wedge \tau_n - S_{k-1} \vee T_{j-1}) + (S_k \wedge \tau_n - T_{j-1})(\alpha-1)e^{(\alpha-1)(B^k_{S_k} - A^j_{S_k})} 1_{\{S_k \le T_{j-1}\}} - (T_j \wedge \tau_n - S_{k-1})(\alpha-1)e^{(\alpha-1)(B^k_{T_j} - A^j_{T_j})} 1_{\{T_j \le S_{k-1}\}} + (S_k - S_k)(\alpha-1)e^{(\alpha-1)(B^k_{S_k} - A^j_{S_k})} 1_{\{T_{j-1} \le S_k \le T_j \wedge \tau_n\}} - (T_j - T_j)(\alpha-1)e^{(\alpha-1)(B^k_{T_j} - A^j_{T_j})} 1_{\{T_{j-1} \le S_k \le T_j \wedge \tau_n\}}\Bigg]\Bigg\}.$$

If we now let $\tau_n \to \infty$, we have

$$= E_{Q^V}\left\{\sum_{j=1}^{N^T}\sum_{k=1}^{N^S}\Big[(S_k \wedge T_j - S_{k-1} \vee T_{j-1}) - |S_k - T_{j-1}|(\alpha-1)e^{(\alpha-1)|B^k_{S_k} - A^j_{S_k}|}1_{\{S_k \le T_{j-1}\}} + |T_j - S_{k-1}|(\alpha-1)e^{-(\alpha-1)|B^k_{T_j} - A^j_{T_j}|}1_{\{T_j \le S_{k-1}\}} + 0 - 0\Big]\right\}.$$

Also, by the same reason, we have

$$E_{Q^V}\left[\int_0^\tau e^{(\alpha-1)(B_t - A_t)}dt\right] = E\left\{\sum_{j=1}^{N^T}\sum_{k=1}^{N^S}\Big[(S_k \wedge T_j - S_{k-1} \vee T_{j-1}) - |T_j - S_{k-1}|(\alpha-1)e^{(\alpha-1)|A^j_{T_j} - B^k_{T_j}|}1_{\{T_j \le S_{k-1}\}} + |S_k - T_{j-1}|(\alpha-1)e^{-(\alpha-1)|A^j_{S_k} - B^k_{S_k}|}1_{\{S_k \le T_{j-1}\}} + 0 - 0\Big]\right\}.$$

Hence

$$CRCRIG_\alpha(N^T, N^S) = E_{Q^V}\left\{\sum_{j=1}^{N^T}\sum_{k=1}^{N^S}\Big[2 \times (S_k - T_{j-1}) + (T_j - S_{k-1}) + |T_{j-1} - S_k|(\alpha-1)[e^{-(\alpha-1)|A^j_{T_j\wedge S_k} - B^k_{T_j \wedge S_k}|} - e^{(\alpha-1)|B^k_{T_j\wedge S_k} - A^j_{T_j\wedge S_k}|}]1_{\{S_k \le T_{j-1}\}} + |S_{k-1} - T_j|(\alpha-1)[e^{-(\alpha-1)|B^k_{T_j\wedge S_k} - A^j_{T_j\wedge S_k}|} - e^{(\alpha-1)|A^j_{T_j\wedge S_k} - B^k_{T_j\wedge S_k}|}]1_{\{T_j \le S_{k-1}\}}\Big]\right\}.$$

and as τ_n goes to ∞ we have

$$CRCRIG_\alpha(N^T, N^S) = E_{Q^V}\left\{\sum_{n=1}^{\infty}\Big[|V_n - V_{n-1}| + (\alpha-1)|V_n - V_{n-1}|(e^{-(\alpha-1)|A^n_{V_n} - B^n_{V_n}|} - (e^{(\alpha-1)|A^n_{V_n} - B^n_{V_n}|})\Big]\right\},$$

as required. □

Acknowledgements The authors express their sincere thanks to the Editor, Associated Editor and the reviewers for their usefull comments and suggestions on an earlier version of this manuscript.

Author Contributions Both authors prepared and reviwed the manuscript.

Funding No funding are received to assist with the preparation of this manuscript

Data Availability No datasets were generated or analysed during the current study.

Declarations

Competing Interests The authors declare no competing interests.

Financial Interests The authors declare they have no financial interests.

References

Arjas E, Yashin A (1988) A note on random intensities and conditional survival functions. J Appl Probab 25:630–635

Aven T, Jensen U (1999) Stochastic Models in Reliability. Springer, New York

Barrapour S, Rad AH (2012) Testing goodness of fit for exponential distribution based on cumulative residual entropy. Comm Stat-Theory Methods 41:1387–1396

Bremaud P (1981) Point Processes and Queues: Martingales Dynamics. Springer, New York

Bueno VC, Balakrishnan N (2022) A cumulative residual inaccuracy measure for coherent systems at component level and under nonhomogeneous Poisson processes. Probab Eng Sci 36:294–319

Bueno VC, Balakrishnan N (2024) An inaccuracy measure between non-explosive point processes with applications to Markov chains. Adv Appl Probab 56:735–756

Capaldo M, Di Crescenzo A, Meoli A (2024) Cumulative information generating function ad generilized Gini function. Metrika 87:775–803

Clark DE (2020) Local entropy statistics for point processes. IEEE Trans Inf Theory 66(2):1155–1163

Di Crescenzo A, Longobardi M (2015) Some properties and applications of cumulative Kullback-Leibler information. Appl Stoch Models Bus In 31(6):875–891

Golomb S (1966) The information generating function of a probability distribution. IEEE Trans Inf Theory 12:75–77

Guiasu S, Reischer C (1985) The relative information generating function. Inform Sci 35:235–241

Jeffreys H (1946) An invariante form for the prior probability in estimations problems. Proc Royal Soc Séries A 186:453–461

Jeffreys H (1948) Theory of Probability. Oxford University Press, Oxford, England

Jacod J (1975) Multivariate point processes: predictable projections, Radon-Nikodym derivatives, representations of martingales. Z Wahrscheinlichkeit 34:235–253

Kerridge DF (1961) Inaccuracy and inference. J R Statist Soc Ser B 23:184–194

Kharazmi O, Balakrishnan N (2021) Jensen-information generating function and its connections to some well-known information measures. Stat Probab Lett 170:108995

Kharazmi O, Balakrishnan N (2023) Cumulative and relative cumulative residual information generating measures and associated properties. Comm Stat Theory Methods 52:5260–5273

Kwiecinski A, Szekli R (1991) Compensator conditions for stochastic ordering of point processes. J Appl Probab 28:751–761

Kullback S, Leibler RA (1951) On information and sufficiency. Ann Math Stat 22:79–86

Kullback S (1968) Information Theory and Statistics. Dover Publishing, New York

Kumar V, Taneja HC (2012) On dynamic cumulative inaccuracy measure. Proceedings of the world congress on engeneering, July 4 - 6, London, England

Kumar V, Taneja HC (2015) Dynamic cumulative residual and past inaccuracy measures. J Statist Theory Appl 14:399–412

Norros I (1986) A compensator representation of multivariate life length distributions, with applications. Scan J Statist 13:99–112

Park S, Rao M, Shin DW (2012) On cumulative residual Kullback - Leibler information. Stat Probab Lett 82:2025–2032

Rao M, Chen Y, Vemuri BC, Wang F (2004) Cumulative residual entropy: a new measure of information. IEEE Trans Inf Theory 50:1220–1228

Rao M (2005) More on a new concept of entropy and information. J Theor Probab 18:967–981

Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27:379–423

Zamani Z, Kharazmi O, Balakrishnan N (2022) Information generating function of record values. Math Methods Statist 31:120–133