Measuring uncertainty, transfer entropy and G-causality In Environmental Economics

Halkos, George and Kitsos, Christos

University of Thessaly, University of West Attica

19 August 2024

Measuring uncertainty, transfer entropy and G-causality In Environmental Economics

George Halkos & **Christos P. Kitsos**

University of Thessaly   University of West Attica

Abstract

Human activities have created environmental degradation with the internalization of the resulting externalities having been the main concern for policy makers worldwide. Uncertainty and unconvincing scientific evidence of various biophysical processes are present in many planned environmental policies. An important source of model uncertainty is accounted by entropy with the typical normal distribution being inadequate in such analyses challenging for more sensible approximations. The problems are, from one side the fat tails characteristic in this area and on the other side what probability density function (pdf) to be chosen. The choice of the appropriate probability model describing the phenomenon that is the pdf is a main priority in any decision making planning. Here we pay attention on the entropy and even more on the transfer entropy in Environmental Economics and the existing underlying uncertainty based on the probability theory. We show that the γ-order Generalized Normal distribution covers both requests, due the "International constant" (γ/(γ-1))**(γ/(γ-1)), leading to a number of pdf, and the Logarithm Sobolev Inequalities (LSI), which provide a solid background.

Keywords: Uncertainty; Environmental Economics; transfer entropy; G-causality.

JEL Codes: Q50; Q56; Q58; C00; C46.

1. Introduction

Uncertainty and entropy are two terms first adopted in the development of Thermodynamics with the pioneering work of Schrodinger (1946). Afterward it was adopted in Physics, then in Mathematics and Statistics and now are applied in various fields, covering Thermodynamics, with the pioneering work of Landau and Lifshitz (1959) followed by Mandl (1988) among others. An uncertain situation means not known beyond doubt, not having complete knowledge. So it is clear that uncertainty is not the error, but it is somehow related with it. In principle, uncertainty is the lack of certainty and as far as the Environmental Economics is concerned, Halkos and Kitsos (2018a) tackled the problem investigating a number of possible causes of uncertainty providing possible measurements of it.

At the early stages of developing such a theoretical approach were not all satisfied. Typical example is Caratheodory's (1909) axiomatic formulation of the Second law of Thermodynamics, which is considered as one of the standard forms of formulation of the law. However, it was oriented more as advanced Mathematics than Physics. The formulation was strongly criticized by Max Planck. We avoid considering such problems from other research fields, but still each problem needs its own Mathematical foundation, especially Physics (Kitsos, 2015). That is there are always those who do not wish to "enlarge" the existed believe, even when the new framework is solid, close to the existent new developments.

Entropy describes the disorder of a system, the "change within a closed system". It is based on the Greek prefix "en" implying within and the "trope" related to the root, here meaning the "change". The "closed system" is a request from the physical point of view of entropy. We understand entropy when the temperature of everything in the room evens out. Entropy is central to the second law of thermodynamics, known as Caratheodory's theorem (among others, Zachananoglou, 1973).

In principle and in mathematical terms, uncertainty links a real-valued function of events in a probability space, which depends on the probability law of the events under consideration. Moreover, events with probability one have zero uncertainty, as we are “certain” that it will take place. Besides as the probability of the event drops, the uncertainty of the event increases. It is then asked to resulting, for the independent events, the uncertainty of the occurrence of two of them - the events under study - to be such as the sum of their individual uncertainties. Looking for a measurable function, reflecting these imposed important requirements, we can conclude that the information of event E, I(E), linked with its probability, $P(E)$, obeys to

$$I(E) = -clog\,[P(E)] \tag{1}$$

with c>0 a given constant.

In this paper we try to apply these ideas, providing more emphasis on the entropy and particularly to *transfer entropy* (TE), working on a family of distributions, to be used in various environmental problems. We have already worked with γ-order generalized normal $N_\gamma(\mu,\Sigma)$ (Halkos and Kitsos, 2018b), introduced by Kitsos and Tavoularis (2009), and it has been already adopted for an extension to TE under the $N_\gamma(\mu,\Sigma)$ (Hlavackova - Schlinder, 2011). We shall try to extend such a work to Environmental Economics working with $N_\gamma(\mu,\Sigma)$ and the extensions on Transfer Entropy (TE).

The structure of the paper is as follows. In section 2 we formalize the existent background, while in section 3 we discuss the appropriate points for $N_\gamma(\mu,\Sigma)$. Section 4 is devoted to the extension that $N_\gamma(\mu,\Sigma)$ can offer to entropy and uncertainty. In section 5 we introduced the transfer entropy especially to Environmental Economics (TEEE). The last section concludes the paper.

2. Background

Uncertainty and entropy, as well as the fundamental approaches in Quantum Statistical Mechanics share a solid background from Physics with Statistics. In 1865, Rudolf Clausius (1822-1888) introduced the term entropy, using Greek terminology as the energy has still the same orientation as etymologists believe. Around 1875, Ludwing Bollzmann (1844-1906) put entropy into the probabilistic framework and developed the Statistical explanation of the Second Law of Thermodynamics. But the fundamental push to probability was given by Shannon (1948), working on formation theory (Coven and Thomas, 1991) as we will refer to it in section 3.

Environmental concerns have been considered as crucial and substantial part of assessing energy conversion systems. The thermodynamic approach implies a profound evaluation of the notion of sustainability as the second law of Thermodynamics counteracts the ability of any open and developing system to preserve itself sustainably without rewarding itself of an uninterrupted supply of low entropy that is high specific exergy input. Exergy in the form of available energy is a crucial notion in thermodynamics. If a society is considered as an open system, its ability to develop in a sustainable manner is a function not only on how it uses the conventional energy sources but also on the speed it exploits the renewable energy sources (Sciubba, 2021).

In economic theory and not only, a number of uncertainty sources, such as model choice uncertainty, data uncertainty, the right mathematical specifications chosen might cause uncertainty as extensively discussed in Halkos and Kitsos (2018a, b). The three different research areas: Risk Analysis, Uncertainty and Entropy are supporting different lines of thought, through Statistics and are not identical. Having different initial points, Risk Analysis have arisen from cancer and now is involved in many areas, such as Environmental Risk; while the other two initiated by Physics, improved by Statistics and we try to prove how important are in Environmental Economics, adopting the Statistical evolution.

Heizerberg's uncertainty principle, in the first introduction in 1927 (Heizerberg, 1927) as a part of the mathematical framework of quantum physics stated that: It is not, in general, possible to predict the value of a quantity with arbitrary certainty, even if all the initial conditions are well imposed and specified. Under uncertainty the more precisely the position Δx of some particle is determined, the less precisely its momentum Δp can be predicted and concluded that: $\Delta_X \Delta_P \approx h$ with h being the full planck constant. Later, Weyl (1928) proved that**:**

$$\sigma_X \sigma_P \geq \frac{h}{2} \; , \; h = \frac{h}{2\Pi} \tag{2}$$

With h' the reduced planck constant h.

There are cases where the information needs to be evaluated under the line of thought as in (1). In such a case for a measurable partition, τ, of the space information function, I(τ), can be considered as:

$$I(\tau) = (-c) \sum_{E \in \tau} \log[P(E)] X_E$$

With X_E being the characteristic function and c is chosen to yield logarithm to base 2, while the expected value of $I(\tau)$ provides the entropy of the partition.

It was Shannon (1948) who thought to provide the information in units of bits. Therefore, he adopted the base-2 measure of entropy, using the logarithm of base 2, in such a way that for given random variable (rv) X, with probability density function $p(x)$ then for the discrete case the entropy of X, $H(X)$, is:

$$H(X) = -\sum_{\chi\in\mathbb{R}} p(x)\log p(x) = E[-\log p(x)] \qquad (3)$$

with E[.] denoting, as usually, the expected value.

Under the same line of thought the conditional entropy of the rv Y given the rv X is defined as:

$$\text{H(Y|X)} = -\sum_{\chi,y\in\mathbb{R}} p(x)\log p(y|x) \qquad (4)$$

While the joined entropy can be, briefly, defined as:

$$\text{H(X, Y)} = \text{H(Y|X)} + \text{H(X)} \qquad (5)$$

When the comparison of two distributions is requested, the Kullback–Leibler divergence, D_{KL}, is adopted, provided that $p(x)$ is the "true" distribution probability and the $q(x)$ probability is asked to be evaluated "how far" is from $p(x)$ then:

$$D(p,q) := D_{KL}(p(x) \parallel q(x)) =$$

$$= \sum_{x\in X} p(x)\log\frac{p(x)}{q(x)} = \sum_{x\in X} p(x)\log p(x) - \sum_{x\in X} p(x)\log q(x) =$$

$$= -\sum_{x\in X}[-p(x)\log p(x)] - \sum_{x\in X} p(x)\log q(x) =$$

$$= \text{E[-log(q(x))]–E[-log(p(x))]} = H_q(X) - H_p(X) \qquad (6)$$

The first line is the notation, the second line is the definition, the third line is a trivial new presentation, which with the last line we can see that the definition is the subtraction of the entropy of X due to "valid" p(x), $H_p(X)$ from the entropy due to "examined" q(x), $H_q(X)$.

In continuous case, the summation is replaced by the integration and is usually referred as "differential entropy". Notice that the Kullback – Leibler divergence, $D = D_{KL}$ as in (6), with a number of notations, with the compact presentation $D = D(p,q)$, will he proved useful in the sequence, it is not symmetric. Actually, the distance measure in Statistics is not exactly what we mean as "distance" in Linear Algebra.

But the lack of symmetry does not create any problem in Environmental Studies. We usually observe the source point of pollution, say at the origin $O(0,0,0)$, and the pollution is moved to point $P_1(X_1, Y_1, Z_1)$ in (Euclidean) distance $OP_1 = L_1$. We are working assuming continuous distribution functions (df) to approach the pollution at every point of investigation. At point O the distribution function of the initial pollution, we say that it is f_0, while at point P_1 is f_1, therefore the K-L divergence is:

$$D_{01} = D_{01}(f_0, f_1)$$

and can be evaluated. Following the same direction, a sequence of points $P_i(X_i, Y_i, Z_i)$, $i = 1,2,\dots,n$ is selected.

Therefore, we can evaluate the K-L divergence between the selected points, and each one from the origin i.e.

$$D_{oi} = D_{oi}(f_0, f_i) \;, i = 1, 2, \dots, n-1$$

$$D_{ij} = D_{ij}\left(f_i, f_j\right) \;,\; j \neq i \,.$$

We can say that the influence of the source of the pollution vanishes when

$$f_i \approx f_{i+1} \Leftrightarrow f_i(x) = f_{i+1}(x)$$

The two successive points $P_i(X_i, Y_i, Z_i)$ and $P_{i+1}(X_{i+1}, Y_{i+1}, Z_{i+1})$ provide the same **df** and thus the K-L divergence is zero, practically very small as:

$$D_{oi}(f_o, f_i) = D_{oi+1}(fi, f_{i+1})$$

Equivalently $|D_{oi} - D_{oi+1}| < \varepsilon$ i=1,2,..,n-q ε>0. (7)

and $\left|D_{ij} - D_{ij+1}\right| < \varepsilon'$ j=i+1ε’ > 0.

The evaluation of the K-L distance at stage n is crucial, while we can have an information data set of “how far”, in terms of Euclidean distance, we are from the origin. It is clear that the distance from the origin, depends on the pollution measurements the investigator collects. If we keep the measurements every d, say, Euclidean distance, the total distance, from the origin, at the end is d_T = (n-1) d.

It would be desirable to work with a specific "large family" of distributions to keep the evaluation of distribution easier. We decided to work with the family of the γ-order Generalized Normal distributions, with shape parameter γ

$$G_\gamma = \{N_\gamma(\mu, \Sigma), \gamma \in \mathbb{R} - [0,1]\} \quad (8)$$

We introduce the definition and probability characteristics of this family in the next section, while section 4 is devoted to the extension of uncertainty and entropy measures it can obtained for the family G_γ.

3. γ-ordered Generalized Normal N_γ (μ, Σ)

The multivariate normal distribution (p variables) is well known, with a large number of applications, in various fields of interest (Anderson, 2004) with Econometrics being one (Halkos, 2019 among others). There are some attempts to create a generalized form, by trial and error, working on the coefficients. The only one emerged with an extra parameter, γ, from a Logarithm Sobolev Inequality (LSI) by Kitsos and Tavoularis (2009), provided an extra shape parameter, and a strong theoretical background for its existence. These were a number of technicalities to generalized the well-known $N(\mu, I\sigma^2)$, the p-variate normal with the vector $\mu \in \mathbb{R}^P$ and covariance matrix $\Sigma \in \mathbb{R}^{Px_P}$.

The γ-ordered generalized normal with an extra shape parameter

$$\mathbb{R} - [0,1] \;\rightarrow\; \gamma \;\in\; \mathbb{R} - [0,1]\,,$$

will be denoted by $N_\gamma(\mu, \Sigma)$ and density function $f_\gamma(x)$ equals to:

$$f_\gamma(x) = C_P[\det(\Sigma)]^{-\frac{1}{2}} exp\left\{-\frac{\gamma-1}{\gamma}[Q(x)]^{\frac{\gamma}{2(\gamma-1)}}\right\} \quad (9)$$

with $x, \mu \in \mathbb{R}^P$ and

$$Q(x) = (x-\mu)\Sigma^{-1}(x-\mu)^T$$

With the transpose $a^T \in \mathbb{R}^{px_1}$, when $a \in \mathbb{R}^{1x_P}$

$$C_P = \pi^{-\frac{p}{2}} \frac{\Gamma\left(\frac{P}{2}+1\right)}{\Gamma\left(p\frac{\gamma}{\gamma-1}+1\right)} \left(\frac{\gamma-1}{\gamma}\right)^{P\frac{\gamma-1}{\gamma}} \tag{10}$$

With $\gamma = 2$ it is obvious that the γ-order generalized normal is reduced to the usual p-variable normal. Moreover $N_\gamma(\mu, \Sigma)$ is a Kotz-type distribution in the sense that

$$N_\gamma(\mu, \Sigma) = Kotz(1, \frac{\gamma-1}{\gamma}, \frac{\gamma}{2(\gamma-1)}; \mu, \Sigma),$$

with Kotz(.) being the appropriate type distribution, and its parameters.

Notice that the density function $f_\gamma(x)$, as in (9), the γ-order generalized normal distribution emerged as an extreme function to Euclidean Logarithm Sobolev Inequality (LSI), and it is not an artificial achievement. This is clear as the term $\gamma_0 = \frac{\gamma-1}{\gamma}$ and its inverse γ_1, moreover the term $(\gamma_1)^{\gamma_1}$ known as “international constant” it is very difficult to be “constructed”. Therefore there is a solid background for the $N_\gamma(\mu, \Sigma)$ (Kitsos and Tavoularis, 2009). The family of distributions has been tackled by Kitsos and Toulias (2011).

Estimation of the parameters, would be a problem, and the MLE, as was considered by Fisher (1925), for the γ-ordered generalized Normal, has been considered $f_\gamma(x)$, the γ-order generalized normal distribution emerged as an extreme function to Euclidean Logarithm Sobolev Inequality (LSI), and it is not an artificial achievement.

The distribution function (9), with normalizing factor (10) forms the family of Generalized Normal distributions as in (8). Therefore different values of γ provide different curves, mainly "fat tailed" distributions.[1] In particular as γ approaches 1 and infinity the $N_\gamma(\mu, \sigma^2 I_P)$ approach the uniform distribution with p=1 and the p-variate uniform equals:

$$U(x) = \frac{\Gamma\left(\frac{p}{2}+1\right)}{\pi^{p/2}\sigma^p} \ , \ |X-\mu| \leq \sigma \tag{11}$$

Moreover as $(\gamma \to \infty)$ the p-variate Laplace is obtained with:

$$L(x) = \frac{\Gamma\left(\frac{p}{2}+1\right)}{\pi^{\frac{p}{2}}\sigma^\rho} \frac{1}{p!} exp\left\{-\frac{|x-\mu|}{\sigma}\right\} \ , \ |X-\mu| \leq \sigma \tag{12}$$

Furthermore, a number of well-known distributions are included in family (8) of the γ-order generalized Normal distributions, defined by (9) and (10), the Dirac distribution being one of them (for details see Kitsos et al., 2012). For these two important distributions in Environmental Economics studies the Kullback–Leibler divergence, for given values of γ can been evaluated. Indeed, the following holds.

Theorem 1: The K-L information between two uniforms with $\sigma_1 \geq \sigma_2$ and two Laplace as in (11) and (12) is:

$$D(U_1, U_2) = p \ln \frac{\sigma_1}{\sigma_2} \tag{13}$$

$$D(L_1, L_2) = p\left[\ln \frac{\sigma_1}{\sigma_2} - 1 + \frac{\sigma_1}{\sigma_2}\right] \tag{14}$$

In this section the Kullback–Leibler divergence, for the given introduced γ-order generalized Normal, for different values of γ have been evaluated. Next step is the introduction of the appropriate Entropy measures for the distribution function (9), generator of the family of distributions as in (8).

---

[1] To work with "fat tailed" or "heavy tailed" distributions see Furioli, et.al. (2022), who although they are adopting Logarithm Sobolev Inequalities (LSI) the distributions discussed are rather complicated to be applied in practical problems, such as the Cauchy distribution**.**

4. Entropy and uncertainty for $N_\gamma(\mu, \Sigma)$

The Shannon Entropy, H(.), of the random variable $X \sim N_\gamma(\mu, \Sigma)$, (Kitsos, Tavoularis, 2009) is:

$$H_\gamma(X) = p\frac{\gamma-1}{\gamma} + \log\frac{(\det(\Sigma))^{\frac{1}{2}}}{Cp} C_P \ as\ in\ (10) \quad (15)$$

Depending on γ every member of the family G_γ obtains the Entropy measure, which in case of γ=z coincide with the p-variate normal $N(\mu, \Sigma)$:

$$H_2(X) = \frac{1}{2}\log\{(2\pi e)^p |\det(\Sigma)|\} \quad (16)$$

With γ=1, the Laplace case the entropy for the multivariate Laplace is:

$$H_L(X) = p + \log\frac{p!\,\pi^{\frac{p}{2}}}{\Gamma\left(\frac{p}{2}+1\right)}[\det(\Sigma)]^{\frac{1}{2}} \quad (17)$$

When p=1 for the uniform distribution $U(\mu-\sigma, \mu+\sigma)$, for the normal, $N(\mu, \sigma^2)$ and for the Laplace $L(\mu, \sigma)$ it can be evaluated respectively:

$$h_U(X) = \log 2\sigma$$

$$h_N(X) = \log(2\pi e\sigma)^{\frac{1}{2}} \quad (18)$$

$$h_L(X) = 1 + \log 2\sigma$$

Proposition 1: It holds

$$h_u^{(x)} < h_N^{(x)} < h_V^{(x)} < h_L^{(x)} \quad (19)$$

Proof:

Indeed with σ> 0 it holds

$$2\sigma < 2\pi\sigma < 2\pi e\sigma$$

thus:$h_u(X) = \log(2\sigma) < \log(2\pi e\sigma) < \frac{1}{2}\log(2\pi e\sigma) = h_N$

and: $h_u(X) = \log(2\sigma) < 1 + \log(2\sigma) = h_L$

From (16) it is easy to be verified that the more the involving parameters p, the larger the (differential) entropy is, i.e. there is an order of the entropy depending on the number of the involving parameters. This implies that in Environmental Economics problems we are working with the accurate number of parameters, although this is another problem:

$$H_2^P(X) \leq H_2^{p+1}(X) \leq \tag{20}$$

The following Theorem is due to Kitsos and Toulias (2010) information, and provides the evaluation of the K-L divergence between generalized Normal, from the defined family as in (8).

Theorem 2: Let two γ-order generalized normal distributions $N_\gamma(\mu_1, \sigma_1^2 I_P)$ and $N_\gamma(\mu_2, \sigma_2^2 I_P)$, with density functions f_1, f_2 respectively, from the family G_γ. Their K-L information is:

$$D_\gamma(f_1, f_2) = C_p \frac{1}{\sigma_1^P}\left\{\frac{p}{2}\left[\log\left(\frac{\sigma_2^2}{\sigma_1^2}\right) I_1 - I_2 + I_3\right]\right\} \tag{21}$$

With the integrals I_i ,i = 1, 2, 3 defined as:

$$I_1 = \int_{\mathbb{R}^P} \exp[-q_1(x)]d_x, \qquad with \;\; q_i(x) = \frac{\gamma-1}{\gamma}\left[\frac{1}{\sigma_1} \| x - \mu_i \|\right]^{\frac{1}{\gamma-1}}, \;\; i = 1,2$$

and

$$I_2 = \int_{\mathbb{R}^P} \exp[-q_1(x)]q_1(x)d_x$$

$$I_3 = \int_{\mathbb{R}^P} \exp[-q_1(x)]q_2(x)d_x$$

Notice that ‖.‖ represents the appropriate norm.It is clear that the entropy is a function only of the standard deviation (the variance in principle).

5. Transfer entropy and G-causality

Uncertainty is related to entropy, which is associated with entropy type Fisher's information, describing different information measures. When we refer to observed data and possible causalities between them, we refer to G-causality (G-C) in honors of Granger (1969) pioneering work and his Nobel Prize Lecture Granger (2003).

There is obviously uncertainty to be measured: what is caused from a process as the new result. Therefore we are looking for a "causality measure" and possible bounds for it. Since the time of Heisenberg (1927) the bounds on the uncertainty are investigated.

As such measure of uncertainty in G-C we chose the entropy type of information so popular in Cryptography. It is necessary to give "more possibility" to the tails, that is adopted the N_γ (μ, $I\sigma^2$), introduced in section 3. The γ-order normal, $N_\gamma(\mu, I\sigma^2)$ as above, offers fat tails and an accepted interpretation for the application, as it is a normal distribution with an extra "shape parameter γ". This is why it was adopted working on the equivalence of G-C and Transfer Entropy (TE), (Hlavackowa –Schlinder, 2011).

In principle the transfer entropy of Y to X given W is defined due to the "history" of the event (Hlavoskova–Schlinder, 2011; Hlavoskova–Schlinder et al. (2016) as:

$$TE\{Y \to X \mid W\} = H(X \mid X^- W^-) - H(X \mid X^- \oplus Y^- \oplus W^-) \tag{22}$$

where the "-" is also referred to "the history" or to "past". Recall the definition of entropy H(.) and the conditional entropy, H(.|), see (4), as well as the definition $\oplus$, Barnet *et al.* (2009) proved that, in a compact presentation, the TE for the G-C is reduced to GTE and holds:

$$\mathrm{GTE}\{Y \to X \mid W\} = \ln[\Theta(X \mid X^- \oplus W^-)] - \ln[\Theta(X \mid X^- \oplus Y^- \oplus W^-)] \tag{23}$$

With the nxn covariance matrix, and the nxm matrix of cross covariance defined as:

$$\mathrm{Gov}[X_i, X_j] = \Theta(X) \in \mathbb{R}^{nxn}$$

$$\mathrm{Gov}[X_i, Y_j] = \Theta(X, Y) \in \mathbb{R}^{nxm}$$

Respectively, for the given jointly distributed multivariate vectors

$$X = (X_1, \dots, X_n) \in \mathbb{R}^n \ , \qquad (Y_1, \dots, Y_m) \in \mathbb{R}^m \ ,$$

$$X\oplus Y = (X_1, \dots, X_n, Y_1, \dots, Y_n) \in \mathbb{R}^{(n+m)}$$

Now, given $\Theta(X)$ the covariance matrix of random variable $X \in \mathbb{R}^n$ and $\Theta(X, \Upsilon)) \in \mathbb{R}^{nxm}$ the cross covariance of the random vectors X and Y with $m \neq n$, we define:

$$\Theta(X|Y) = \Theta(X) - \Theta(X, \Upsilon)[\Theta(\Upsilon)]^{-1}\Theta(X, \Upsilon)^T \tag{24}$$

under the assumption that the inverse of $\Theta(\Upsilon)$ exists in $\mathbb{R}^{mxm}$. Notice that $\Theta(X|Y)\epsilon R\ nxn$ as both parts at the right hand side of (24) as in nxn. It is easy to see, with the formulation we offer, the similarity of (22) and (23).

According to Granger (2003) a consequence of this attempt is that the casual variables can help to forecast the effect variable after other data has been first used. Although this statement is too optimistic, still we have more information for the "future stage" being at the "current stage".

Therefore the extension to Transfer Entropy problem for the γ-order Normal distribution $N_\gamma(\mu\, ; \Sigma)$ was necessary and Hlavackova–Schlinder et al. (2016) provided a number of results for it. The important result, for our development is the provided in the Appendix A Theorem, which evaluates, although in a complicated way the fact that

$$TE_\gamma\{W \to U \mid V\} \tag{25}$$

exists and it is evaluated.

The crucial result from this Theorem, in Appendix A, is that it generates the results for the classical Normal distribution, which are provided for the case of γ = 2: the transfer entropy with normally distributed and spherically contoured joint distribution vanishes.

Proposition 2: In the case of the normal distribution $N(\mu, \sigma^2 I_q)$:

$$TE_2\{W \to U \mid V\} = 0 \tag{26}$$

Recall that the conditional mutual information I(U,W|V) vanishes, if and only is (iff), the conditional random variables U|V and W|V are independent. Therefore transfer entropy can be expressed also by means of conditional mutual information

$$TE_2\{W \rightarrow U \mid V\} = \mathrm{I}\,(\mathrm{U}, \mathrm{W}|\mathrm{V}) \qquad (27)$$

This is a crucial result as the transfer entropy with normally distributed variable with a spherically contoured joint distribution is zero. So we need more general forms, generalized multivariate normal $N_\gamma(\mu, \sigma^2)\,, \gamma \neq 2$, elliptically contoured normal applied in Theorem of Appendix A.

Proposition 3: The Transfer Entropy

$$TE\gamma\{W \rightarrow U \mid V\} \text{ with } U \oplus V \oplus W \sim \text{Laplace} L^q(\mu, \sigma^2, I_2)\,,$$

$$i.e.\ when\, \gamma \rightarrow \pm\infty, with\ \ q = 1 + m + n$$

exists and can be evaluated.

Proposition 4: Under certain conditions for given pdf p, q, namely p is the joint pdf. of X and Y and p exists for stationary random processes Y and X, depending on the estimated value of x, up to time t=i. given the "history" of x and y i.e. the pdf

$$p=p(x_{est}|\, x^-, y^-)) \quad \text{with } p \sim N_\gamma(\mu_0, \sigma^2 I_r)$$

and the pdf q is

$$q = q(y) = p(x_{est}, y^-)\, p(x_{est}, x^-)/p(x^-) \ \text{ with } q \sim N_\gamma(\mu_0, \sigma_1^2 I_r)$$

Then: $$TE\{Y \rightarrow X\} = D_{KL}(p \parallel q) = r\left\{\log\left(\frac{\sigma_0}{\sigma_1}\right) - \frac{\gamma-1}{\gamma}\left[1 - \left(\frac{\sigma_1}{\sigma_0}\right)^{\frac{\gamma}{\gamma-1}}\right]\right\} \qquad (28)$$

For details in the case of $\sigma_0 \neq \sigma_1$, see Kitsos and Toulias (2010).

Following Jeffreys (1946) recipe for Symmetric Transfer Entropy (STE) we can define it as: $STE\{Y \rightarrow X\} = 1/2\{\, D_{KL}(p \parallel q) + D_{KL}(q \parallel p)\}$

Due to this essential point the (28) can be reduce to a symmetric one as:

$$STE\{Y \rightarrow X\} = r/2\left\{\left[2 - \left(\frac{\sigma_1}{\sigma_0}\right)^{\frac{\gamma}{\gamma-1}} - \left(\frac{\sigma_1}{\sigma_0}\right)^{\frac{-2\gamma}{\gamma-1}}\right]\right\} \qquad (29)$$

This is an easy relation to be calculated, to obtain symmetric transfer entropy in cases were symmetry is essential in Environmental Economics.

From the above mentioned useful theoretical results the following points must be considered with a special attention:

i. Keeping γ stable, then

$$D_{KL}^{r} < D_{KL}^{r+1}.\ \pi = 1, 2, 3\ldots \tag{30}$$

Due to (30) the Laplace distribution, provides the lower bound, when the shape parameter tends to infinity

ii. For given r=m+n+1, $\mu_1 = \mu_0$ and for $\gamma_1 < \gamma_2$ then:

$$D_{KL,\gamma 1} > D_{KL,\gamma_2} \tag{31}$$

iii. Notice that the value of the shape parameter participates with a special weight to the analysis we would like to proceed. For value γ=2 it vanishes! Therefore the classical Normal distribution is inadequate for such applications. A rough explanation can be justified as "there is any disorder between Normals", that is we need a more realistic approach – the fat tailed family of distributions as the family (8) with members as in (9).

iv. The exact calculations through given states of the process is difficult. To obtain $TE\{\boldsymbol{X} \rightarrow \boldsymbol{Y}\}$ is simple through the existent results, especially (28), a quite easy to be calculated way of the involved transfer entropy.

6. Discussion

This paper is based on our believes that quantified methods and optimization procedures need the Mathematical background, and a common understanding and link with the area of interest. Our main concern here is in Environmental Economics in which there is no doubt that an underlying uncertainty exists based on the probability theory.

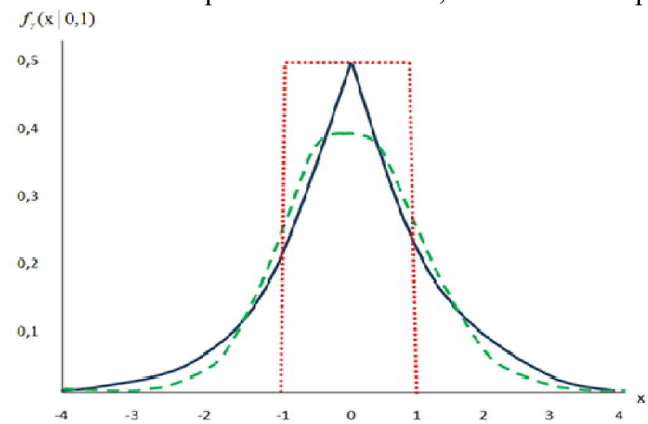
An essential source of model uncertainty measured by entropy comes into view from the imposed assumptions and the assumed distributions to be considered. We are referring to entropy as it is strongly related to a system's disorder and uncertainty and is related to variance. As it is applied to a lot of engineering problems we believe that this idea can be adapted in Economics in addition to a measure of existing uncertainty.[2]

It is clear that the entropy depends only on the variance-covariance matrix Σ or σ in case $p=1$. In practice this means that the uncertainty is irrelevant to mean value μ (of the pollution centre to an industry, say) but depends on the standard deviation (the experimental error), that is how much is expanding from the centre. The Uniform distribution can be adopted if it is assumed that pollution levels are (almost) the same around the area [a, b], while the Laplace when it is assumed a "sharp explosion" around the center and much lower far from it. Estimates of (9) can be obtained in practical situations. Figure 1 is a graphical representation of the relationship between Uniform, Normal and Laplace in the univariate case.

In simple words, this implies that the choice of the appropriate probability model describing the phenomenon that is the probability density function (pdf) is a main priority in any decision making planning. The typical normal distribution is inadequate in such cases demanding more realistic approximations with the fat tailed family of distributions as presented here to facilitate in these directions.

[2] The researcher, in principle, may not be a Mathematician, so no need to know the existing development, but the solid results and the computational applications (Kitsos and Nyamsi, 2024). Exploring the economy–environment relationship in the case of air pollution (Halkos, 2013) but also the effects of government expenditure on the environment (Halkos & Paizanos, 2017), coping with energy poverty (Halkos and Gkampoura, 2021) and exergy (Sciubba, 2021) will be helped among others by the proposed analysis.

Figure 1: Relationship between Uniform, Normal and Laplace graphically



That is to say, a problem associated with the normal distribution is the "fat tails" and there are situations where the assumed normal distribution in tails "contains more probability" than the usual 0.05. This is spot on in a number of economic applications and surely in various environmental problems where pollution has an effect on the "tails" more than 0.05. This implies a urgent need to acquire a distribution with "fat tails" and consequently we are looking for a class of distributions, were normal (the "ideal situation") might be a member of this class requiring a generalization of normal. In these lines, the γ-order Generalized Normal Distribution (γ-GND) emerged from a completely mathematical problem with Logarithm Sobolev Inequality (LSI) providing a solid background for it.

Our target it is not to clarify a number of fundamental ideas in other fields, but to improve their understanding, and applicability in Environmental Economics, were a very useful tool is in hands: the γ-ordrer Generalized Normal distribution. To the three main pillars, uncertainly - G-causality - Transfer Entropy, we add Environmental Economics, to obtain new results, to create new theoretical developments, being helpful to the researcher to understand the investigating phenomenon much better.

This will assist in the construction of adequate confidence intervals that will smooth the progress of the planning of effective environmental policies.[3]

Conclusions

In this paper we emphasize on the entropy and more specifically on transfer entropy. We have already shown that assuming the γ-order generalized normal may help extending to transfer entropy. One problem is the fat tails characteristic, in this area, and another problem is what pdf to be chosen. So the family of distributions, the γ- order Generalized Normal distribution covers both requests, due the "International constant" (γ/(γ-1))**(γ/(γ-1)), which leads to a number of pdf, and the Logarithm Sobolev Inequalities (LSI), which provide a solid background.

Therefore uncertainty was extensively discussed through its origins, for the Environmental Economics problem. Moreover G-causality is essential to Economic problems and related to Transfer Entropy definition, highly related with the Information Theory. That is why we tried to clarify and simplify the Physical ideas: to be adopted in Environmental Economics, as are not that complicated any more, and the appropriate packages can offer easy calculations. At the same time, the concepts are essential to the field of our concern: in Environmental Economics.

---

[3] For instance in the case of transfrontier pollution and assuming linearity, the total annual depositions (AD_i) of air pollution (like sulphur) in country i, will be given as (Halkos, 1993, 1994, 1996):

$$AD_i=\Sigma_j d_{ij}(1-\alpha_j)E_j+B_i \qquad \forall i,j \quad i\, , j=1,...,N$$

Where α_i is the abatement efficiency coefficient in country i, d_{ij} is the transfer coefficient from country j to i, indicating what proportions of emissions from any source country is ultimately deposited in any receiving country, B_i is the level of the so-called background deposition attributable to natural sources (such as volcanoes, forest fires, biological decay, etc) in receptor-country i, or to pollution remaining too long in the atmosphere to be tracked by the model, i.e. is probably attributable not only to natural sources but also to emissions whose origin cannot be determined.

APPENDIX A

Transfer Entropy evaluation for the γ-order Generalized Normal distribution

Theorem. The transfer entropy for $U \oplus V \oplus W \sim N_\gamma(\mu, \sigma^2 I_q)$is:

$$TE_\gamma\{W \to U \mid V\} = \log A - B(q,\gamma)S_1(q,\gamma)S_2(q)$$

With

$$A = A(q,m,n,\gamma) = \frac{\Gamma(\frac{q}{2})\Gamma(\frac{m}{2})\Gamma(\frac{1+m}{\gamma_1})\Gamma(\frac{m+n}{\gamma_1})}{\Gamma(\frac{q}{\gamma_1})\Gamma(\frac{m}{\gamma_1})\Gamma(\frac{1+m}{2})\Gamma(\frac{m+n}{2})} \; and \quad \gamma_1 = \frac{\gamma}{\gamma-1}$$

$$B = B(q,\gamma) = \frac{\Gamma(\frac{q}{2})}{2^{\frac{q}{2}}\Gamma\left(\frac{q}{\gamma_1}\right)\gamma_1^{\frac{q+\gamma_1}{\gamma_1}}}$$

$$S_1(q,\gamma) = \sum_{k=0}^{\infty} \frac{\left(\frac{\gamma_1}{2}\right)_k}{k!} \sum_{\lambda=0}^{k} \binom{k}{\lambda} (-1)^\lambda \gamma_1^{\frac{2(k-\lambda)+q}{\gamma_1}} \Gamma(\frac{2(k-\lambda)+q}{\gamma_1})$$

$$S_2(q) = \sum_{I_q^{k-\lambda}}^{\infty} \frac{(k-\lambda)!}{i_1 i_2 \dots i_q} P_q(m,n)$$

with the usual notation valid $(r)_s = r(r-1)\dots(r-s+1)$ and$I_q^{k-\lambda}$ be a set of indices such that

$$I_q^{k-\lambda} = \left\{ in(i_1, i_2, \dots, i_q) \in N^2, \sum_{j=1}^{q} i_j = k-1, i_1 \in (k,\lambda) \;\; or\; t_{m+2} + \dots + t_q \in (k,\lambda) \right\}$$

And $P_q(m,n)$ well defined polynomials depending on $m,n \in N^*$ being or not being both odd or even number, see Hlavackova – Schlinder *et al*. (2016) for details. The P_q (m, n) are well defined polynomials depending on m, n $\in N^*$ being or not being both odd or even numbers, see Hlavackova – Schlinder *et al.* (2016), and their Theorem 6.2 for details.

References

Anderson, T. W. (2004). *An Introduction to Multivariate Statistical Analysis.* 3rd Edition). Wiley India.

Barnett, L., Barett, A.B., and Seth, A.K. (2009). Granger causality and Gaussian variables. Phys. Rev. Lett., **103:** 23.

Caratheodory, C.(1909), The Second Law of Thermodynamics, Benchmark Papers on Energy 5, Edited by Joseph Kestin, DH&R Inc., Stroudsburg, This article was translated expressly for the Benchmark Volume by J Kestin, Brown Univ. from "untersuchungenuber die Grundlagen der Thermodynamik," in Math. Ann. (Berlin), 67, 355-386 (1909).

Cover, T., Thomas, J. (1991). *Elements of Information Theory*. John Wiley & Sons.

Fisher R.A. (1925). Theory of statistical estuation. *Proceedings of the Cambridge Philosophical Society*, **22:** 700-725.

Furioli, G., Pulvirenti, A., Terraneo, E., Toscani, G. (2022). Fokker-Planck equations and one-dimensional functional inequalities for heavy tailed distribution. Milan, J. of Math., 90, 177-208.

Granger, C.W.J. (1969). Investigating causal relations by econometric and cross-spectal methods. *Econometrics*, **37:** 424-430

Granger, C.W.J. (2003). Time series analysis cointegration and applications. Nobe Lecture, Dec. 8, 2003 in: Les Prix Nobel. The Nobel Prizes 2003, Tore Frangsmyn (ed), (Nobel Foundation, Stockholm, 2004) pg. 360-366

Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proc. Roy. Soc A*., **186**::453-461.

Halkos, G.E. (1993). Sulphur abatement policy: Implications of cost differentials, *Energy Policy*, **21(10):** 1035-1043.

Halkos, G.E. (1994). Optimal abatement of sulphur emissions in Europe, *Environmental & Resource Economics*, **4(2):** 127-150.

Halkos, G.E. (1996). Incomplete information in the acid rain game, *Empirica,* **23(2):** 129-148.

Halkos, G. (2013). Exploring the economy–environment relationship in the case of sulphur emissions. *Journal of Environmental Planning and Management* **56(2):** 159-177.

Halkos, G. & Paizanos, E. (2017). The channels of the effect of government expenditure on the environment: Evidence using dynamic panel data. *Journal of Environmental Planning and Management* **60 (1):** 135-157.

Halkos, G. and Gkampoura E.C. (2021). Coping with energy poverty: Measurements, drivers, impacts, and solutions Energies **14(10):** 2807

Halkos, G. (2019). *Econometrics. Theory and Applications*. Disigma Pub. Athens. (in Greek).

Halkos, G. & Kitsos, C.P. (2018). Mathematics vs Statistics in tackling Environmental Economics Uncertainty. MPRA paper, No.: **85280.**

Halkos, G. & Kitsos, C. (2018). Uncertainty in environmental Economics: the problem of entropy and model choice. *Economic Analysis and Policy*, **60:** 127-140.

Heisenberg, W. (1927). Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik* (in German) **43(3)**: 172–198.

Hlavackova – Schlinder, K., Toulias, T.L., and Kitsos, P.C. (2016). Evaluating Transfer Entropy for Normal and γ-order Normal Distribution. Br. I. of Math. & Comp. Science, 17(5), 1-20, Article No.: BJMCS 27377

Hlavackova – Schlinder, K. (2011). Equivalence of Granger Causality and Transfer Entropy: A Generalization. *Appl. Math. Sciences*, **5(73)**: 3637-3648.

Kitsos, C. P. (2015). Technological Mathematics and Statistics., Volume II New Tech Pub. Athens (in Greek).

Kitsos P.C., Tavoularis N. (2009). Logarithm Sobolev Inequalities for information measures. *IEEE Transactions on Information Theory,* **55(6):** 2554-2561.

Kitsos, C. P., Toulias, T. (2010). New Information Measures for the Generalized Normal Distribution. *Information*, **1:** 13-27.

Kitsos, **C.** P ,Toulias, T. L (2011). On the family of the γ-ordered normal distributions. *Far East Journal of Theoretical Statistics,* **35(2):** 95-14,

Kitsos, C. P., T. L. Toulias (2012b). Bounds for the generalized entropy type information measure. *Journal of Communication and Computer*, **9(1)**: 56-64.

Kitsos, C.P. & Nyamsi, U. E. (2024). Applying the Generalized Normal distribution for Modelling Asset Returns Data. **In**: The Fifth International Congress of Applied Statistics (UYIK-2024), Istanbul 21-23 May 2024. Turkey.

Landau, L.D. ,Lifshitz, E.M. (1959). *Statistical Physics*. Pergamon. London

Mandl, F. (1988). *Statistical Physics*, 2d edition. Wliey

Schrodinger, E. (1946). *Statistical Thermodynamics*.Cambridge Un. Press. Cambridge.

Sciubba, E. (2021). A thermodynamic measure of sustainability. *Frontiers in Sustainability,* **2:** 739395.

Shannon, C. (1948). A mathematical theory of communication. *Bell System Tech J.,* **27**: 379-423, 623-656.

Weyl, H. (1928). *Gruppentheorie und Quantenmechanik*. Leipzig:Hizel.

Zachmanoglou, E.C. (1973). Caratheodory's theorem on the second law of thermodynamics. *SLAM. J. Appl. Math*, **25(4):** 592-596.