Annals of the University of Craiova, Mathematics and Computer Science Series
Volume 43(1), 2016, Pages 108–117
ISSN: 1223-6934

*Dedicated to Marius Iosifescu*
*on the occasion of his 80th anniversary*

# Entropy Concepts Applied to Option Pricing

Radu Tunaru

Abstract. Uncertainty is one of the most important concept in financial mathematics applications. In this paper we review some important aspects related to the application of entropy-related concepts to option pricing. The Kullback-Leibler information divergence and the informational energy introduced by Onicescu are the main tools investigated in this paper. We highlight a necessary condition that must be verified when obtaining the probability distribution minimising the Kullback-Leibler information divergence. Deriving a probability distribution by optimising the information energy has some pitfalls that are discussed in this paper.

*2010 Mathematics Subject Classification.* 91G70, 91G20, 62E10.
*Key words and phrases.* Kullback-Leibler divergence measure, informational energy, option pricing, optimisation.

## 1. Motivation

**1.1. Entropy Probabilistic Concepts in Finance.** In probability and statistics, several entropy-related concepts have been successfully defined and applied, starting with the Shannon entropy as described in [Shannon, 1948], then the Kullback-Leibler information divergence measure, introduced by [Kullback and Leibler, 1951] and [Kullback, 1959]. The principle of maximum entropy has been used to define a new family of estimators that can be applied with limited dependent data, when the usual estimation principles like maximum likelihood and least squares fail, see [Csiszar, 1991], [A. Golan and Miller, 1996]. Quite often entropy optimisation methods are designed in conjunction with moment matching conditions. A modern approach to inference problems, including stochastic inverse problems, can be found in [Judge and Mittelhammer, 2012].

In Economics and Finance the entropy related concepts found a natural bedrock of applications and problems for which they offered an elegant solution. Early considerations were described in [Theil, 1967], arguable the first one to draw a line between economics and information theory and [Georgescu-Roegen, 1971], who made a link between economics and physics for environmental economics. [Buchen and Kelly, 1996] and [Stutzer, 1996], who were among the first to derive thew maximum entropy distribution of an asset inferred from option prices.

[Avellaneda et al., 1997], [Avellaneda, 1998] advocated calibrating volatility surfaces using the relative-entropy minimization principle. [Bariviera et al., 2016] used the complexity-entropy causality plane to detect an abnormal movement of LIBOR series around the 2007 crisis.

In a series of papers, [Gulko, 1999a, Gulko, 1999b, Gulko, 2002], Les Gulko reformulated some of the most important ideas related to efficient market hypothesis, bond option pricing and stock option pricing. He showed that Black-Scholes-type formulae reappear under a new set of assumptions based on maximum-entropy formalism.

This paper is structured as follows. In section 2 we describe the modelling set-up for option pricing and the mathematical problem that must be solved in order to derive the probability distribution to be used for option pricing. Our contribution is discussed in section 3. The section 4 contains the main conclusions and some relevant ideas of further research in this area.

## 2. Modelling set-up

**2.1. The Kullback-Leibler Information Divergence.** One of the most famous measures of discrimination of information between two probability measures, or more applied, two probability density functions, is the Kullback-Leibler information divergence defined in [Kullback, 1959] and [Kullback and Leibler, 1951]. An axiomatic derivation of this measure has been proposed by [Tunaru, 1992].

For a continuous random variable $X$ that may be associated with *two* probability distributions $P$ and $Q$, the Kullback-Leibler information divergence is defined by the formula

$$I(P;Q) = \int_{-\infty}^{\infty} p(x) \ln \frac{p(x)}{q(x)} dx \tag{1}$$

where $p$ and $q$ are the probability densities of the probability measures $P$ and $Q$ with respect to the Lebesgue measure[1].

**2.2. The Optimisation Problem.** From a financial mathematics point of view we consider a traded asset with the price process denoted generically by $S$. Since we are mainly concern with one-period modelling we consider that $S_0$ is the value of the asset now[2] at time $t_0$ and we need to evaluate the price of a contingent claim contract paying only at fixed maturity $T > 0$ the payoff $\Psi(S_T)$. We assume that we are in a competitive Arrow-Debreu economy and we consider that there is a sequence of values $S_{t_1}, S_{t_2} \ldots, S_{t_n}$ observed in the past, so $t_n < t_{n-1} < \ldots t_1 < t_0$. Hence, today the analyst has an empirical probability density distribution $g$ constructed from the observed returns series given by $R_i = \ln\left(\frac{S_{t_{i-1}}}{S_{t_i}}\right)$ for all $i \in \{1, \ldots n\}$. We can assume without loss of generality that $t_0 \equiv 0$ and that the frequency interval of observation $t_i - t_{i-1} = T$. This is necessary such that information scaling effects on the probability distributions are captured correctly.

The investor is interested in finding out the probability distribution $f$ of $S_T$ in order to determine the price of the payoff $\Psi(S_T)$ now, at time $t_0 \equiv 0$. Since the party in the financial contract that is selling, i.e. that will have to make the payment $\Psi(S_T)$ at maturity $T$, will need to take into consideration a wide range of possible scenarios, the density $f$ should reflect the maximum degree of uncertainty.

As it is common in finance we assume that we know $S_0$, the volatility parameter $\sigma > 0$ that is the standard deviation of $R_T$ and $p(0,T)$ the price of a zero-coupon

[1] Given the applied nature of this paper we can safely assume that for all practical purposes all probability spaces are endowed with the Lebesgue measure.

[2] Without reduction of complexity we can assume that $t_0 = 0$.

bond price with maturity $T$. Moreover, we shall assume that the market representative agent's belief of the distribution of feasible values of the final asset price $S_T$ is fully characterised by maximum uncertainty.

Thus, here we generalise the entropy pricing theory proposed by [Gulko, 1999b] and we shall assume that the market agent's beliefs on $S_T$ are encapsulated into the probability density $f(R_T)$ with the support of all states given by $\mathcal{D}$, that can be either the entire line $\mathbb{R}$ or a smaller interval. [Gulko, 1999b] uses the Shannon entropy $H(f)$ as a measure of market pricing uncertainty. One criticism that can be brought about this measure is that it ignores completely valuable information from the past about the asset price returns $R$. We would like to identify the probability density $f$ of the future return/state price $S_T$ given the current knowledge of the empirical distribution $g(R_T)$. Our proposed solution is to use the Kullback-Leibler information divergence to define an improved measure of market uncertainty. Therefore, extending the entropy option pricing principle in [Gulko, 1999b], we propose

**Assumption 2.1.** The market uncertainty index at a future time $T > 0$ about $R_T$ is given by

$$J(f; g) = -\int_{-\infty}^{\infty} f(R_T) \ln \frac{f(R_T)}{g(R_T)} dR_T \tag{2}$$

Clearly $J(f; g) = -I(f; g)$. Since $I(f; g)$ has been used to define the Principle of Minimum Discrimination Information (MDI) which in few words works like this: based on the arrival of new information, a new distribution $f$ is selected such that is the closest possible to the original distribution $g$. In order to maintain a parallel approach with [Gulko, 1999b] we would like to maximize $J(f; g)$ which is equivalent to minimize $I(f; g)$. We are searching for the probability density $f$ that maximizes $J(f; g)$, or equivalently minimize $I(f; g)$, subject to some externally imposed conditions provided by Finance theory. The problem we would like to solve is given in the following proposition.

*Problem*: At time zero assuming that we know $S_0, r, p(0, T), \sigma^2$ the minimum information deviance probability density $f(R_T)$ solves the problem

$$\min_f I(f; g) \tag{3}$$

subject to

$$\int_{-\infty}^{\infty} f(R_T) dR_T = 1 \tag{4}$$

$$f(R_T) > 0 \quad \forall R_T > 0 \tag{5}$$

$$E_f(R_T) = \mu \tag{6}$$

$$var_f(R_T) = \sigma^2 \tag{7}$$

The optimisation problem above has the objective $\int_{-\infty}^{\infty} f(R_T) \ln \frac{f(R_T)}{g(R_T)} dR_T$ which is a convex functional on the convex set of probability densities defined over $(-\infty, \infty)$.

In order to solve this problem one can use use calculus of variations for optimisation of functionals as described in [Luenberger, 1969]. [Friedman et al., 2010] applied the above framework to distribution identification from a large family of skewed generalized $t$ distribution proposed in finance by [Theodossiou, 1998].

## 3. Financial Applications of Informational Measures

### 3.1. Option Pricing by minimisation of the Kullback-Leibler divergence measure.

**Proposition 3.1.** *The distribution minimizing the information divergence given by the Kullback-Leibler functional I(f;g), and that satisfies the conditions*

$$\begin{aligned} \int_{-\infty}^{\infty} f(R_T)dR_T &= 1 && (8)\\ f(R_T) &> 0 \quad \forall R_T && (9)\\ E_f(R_T) &= \mu && (10)\\ E_f(R_T^2) &= \sigma^2 - \mu^2 && (11) \end{aligned}$$

*has a gaussian kernel.*

The proof is based on the calculus of variations for optimisation of functionals described in [Luenberger, 1969]. To this end, as in [Avellaneda, 1998], see also [Borwein et al., 2003], consider the dual problem described by the Lagrangian operator

$$\begin{aligned} \mathcal{L}(f(R_T), \lambda_1, \lambda_2) = I(f;g) + \lambda_1 \left(\mu - \int_{-\infty}^{\infty} R_T f(R_T)dR_T\right) + \\ + \lambda_2 \left(\sigma^2 - \mu^2 - \int_{-\infty}^{\infty} R_T^2 f(R_T)dR_T\right) \\ + \lambda_0 \left(\int_{-\infty}^{\infty} f(R_T)dR_T - 1\right) + \int_{-\infty}^{\infty} f(R_T)\psi(R_T)dR_T \end{aligned}$$

that has the Lagrange multipliers $\lambda_1, \lambda_2, \lambda_0, \psi$. The dual problem is

$$\max_{\lambda,\lambda_0,\psi} \min_{f} \mathcal{L}(f(R_T), \lambda_1, \lambda_2) \tag{12}$$

which is a concave problem in parameters $\lambda_1, \lambda_2$ and $\psi$.

The solution to our optimisation problem is given by those values of parameters $\lambda_1, \lambda_2$ and $\psi$ coming out of the Euler-Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial f} + \frac{d}{dR_T}\left(\frac{\partial \mathcal{L}}{\partial f'}\right) = 0 \tag{13}$$

Remarking that the operator $\mathcal{L}(f, R_T)$ is independent[3] of the derivative $f' = \frac{df(R_T)}{dR_T}$, the Euler-Lagrange equation is

$$\ln f + 1 - \ln g - \lambda_1 R_T - \lambda R_T^2 + \lambda_0 + \psi(R_T) = 0 \tag{14}$$

Using now the Karush-Kuhn-Tucker conditions leads to set $\psi(R_T) = 0$ and after the probability density normalising condition we get the solution

$$f(R_T) = \frac{g(R_T)\exp\left[\lambda_0 - 1 + \lambda_1 R_T + \lambda R_T^2\right]}{\int_{-\infty}^{\infty} g(x)\exp\left[\lambda_0 - 1 + \lambda_1 x + \lambda x^2\right]dx} \tag{15}$$

This solution by design also satisfies the positivity constraint (9) and probability density normalisation (8). Furthermore, since for any fixed $g$ the functional $I(f;g)$ is

[3] Tacitly we are assuming that the probability density $f$ is smooth enough such that $f'$ does exist

a convex functional over the convex set of probability densities defined over the real set $\mathbb{R}$, it follows that the probability density given in (15) is a global optimiser.

The family of distributions described by (15) encompasses the family of gaussian distributions, for example when $g$ is a constant, but it is much more general. While the case of entropy maximization subject to moment constraints has been well-studied in the literature such that we know that if only the first moment is fixed then the distribution that maximizes the entropy is the exponential distribution, and when the first two moments are fixed is the gaussian distribution, one cannot draw similar conclusions when trying to minimize the Kullback-Leibler information divergence measure. The reference probability distribution $g$ is complicating things.

Given the general form arrived at in (15) we can try to prove the existence of a solution and study the uniqueness of this solution. Based on (15) the conditions (8),(10) and (11) are written as

$$H_0(\lambda_0,\lambda_1,\lambda_2) \equiv \int_{-\infty}^{\infty} g(x)e^{\lambda_0-1+\lambda_1 x+\lambda_2 x^2}dx = 1 \tag{16}$$

$$H_1(\lambda_0,\lambda_1,\lambda_2) \equiv \int_{-\infty}^{\infty} xg(x)e^{\lambda_0-1+\lambda_1 x+\lambda_2 x^2}dx = \mu \tag{17}$$

$$H_2(\lambda_0,\lambda_1,\lambda_2) \equiv \int_{-\infty}^{\infty} x^2g(x)e^{\lambda_0-1+\lambda_1 x+\lambda_2 x^2}dx = \sigma^2-\mu^2 \tag{18}$$

For reasons that will become clear immediately we assume that the distribution we are searching for has finite moments up to order four, that is $m_3$ and $m_4$ exists and they are finite. Following an idea from [Calin and Udriste, 2014], denoting by $\Delta$ the determinant $\det\left(\frac{\partial H_i}{\partial \lambda_j}\right)_{i,j\in\{0,1,2\}}$, by Inverse Function Theorem, the system (16) has a unique solution when $\Delta \neq 0$. Observe now that

$$\begin{aligned}
\frac{\partial H_0}{\partial \lambda_0} &= \int_{-\infty}^{\infty} g(x)e^{\lambda_0-1+\lambda_1 x+\lambda_2 x^2}dx = 1 \qquad (19)\\
\frac{\partial H_0}{\partial \lambda_1} &= \int_{-\infty}^{\infty} xg(x)e^{\lambda_0-1+\lambda_1 x+\lambda_2 x^2}dx = \mu\\
\frac{\partial H_0}{\partial \lambda_2} &= \int_{-\infty}^{\infty} x^2g(x)e^{\lambda_0-1+\lambda_1 x+\lambda_2 x^2}dx = \sigma^2-\mu^2\\
\frac{\partial H_1}{\partial \lambda_0} &= \int_{-\infty}^{\infty} xg(x)e^{\lambda_0-1+\lambda_1 x+\lambda_2 x^2}dx = \mu\\
\frac{\partial H_1}{\partial \lambda_1} &= \int_{-\infty}^{\infty} x^2g(x)e^{\lambda_0-1+\lambda_1 x+\lambda_2 x^2}dx = \sigma^2-\mu^2\\
\frac{\partial H_1}{\partial \lambda_2} &= \int_{-\infty}^{\infty} x^3g(x)e^{\lambda_0-1+\lambda_1 x+\lambda_2 x^2}dx = m_3\\
\frac{\partial H_2}{\partial \lambda_0} &= \int_{-\infty}^{\infty} x^2g(x)e^{\lambda_0-1+\lambda_1 x+\lambda_2 x^2}dx = \sigma^2-\mu^2\\
\frac{\partial H_2}{\partial \lambda_1} &= \int_{-\infty}^{\infty} x^3g(x)e^{\lambda_0-1+\lambda_1 x+\lambda_2 x^2}dx = m_3
\end{aligned}$$

$$\frac{\partial H_2}{\partial \lambda_2} = \int_{-\infty}^{\infty} x^4 g(x) e^{\lambda_0 - 1 + \lambda_1 x + \lambda_2 x^2} dx = m_4$$

Thus, the condition $\Delta \neq 0$ is equivalent to the condition

$$(\sigma^2 - \mu^2)[(\sigma^2 - \mu^2)^2 - 2\mu m_3] \neq m_4\sigma^2 - m_3^2 \tag{20}$$

This condition is a necessary condition for the existence and uniqueness of the distribution with minimum Kullback-Leibler divergence and it is less known in the option pricing and calibration financial mathematics literature.

Another important observation regarding the solution (15) is that, when $g \equiv 1$, the solution obtained belongs to the gaussian family. Given that the mean and variance are fixed from conditions (10) and (11), it follows that the Kullback-Leibler uncertainty maximised, or minimum divergence, distribution is $N(\mu, \sigma^2)$. This will automatically imply a lognormal distribution for the underlying variable $S_T$ given by the parameters $\mu$ and $\sigma$. Hence $E(S_T) = \exp\left(\mu + \sigma^2/2\right)$.

Hence, we were able to arrive at a solution describing a Black-Scholes world, *without* the assumptions of a Black-Scholes model, no continuous-time model for the underlying asset, no requirement of a complete market [Bjork, 2009]. This approach may prove very useful in the context of real options where the underlying asset is only observable but no tradeable. For example, if $S_T$ is the value of a real-estate index, one can observe the values for $S_T$ but will not be able to buy and sell that index. Similarly, $S_T$ could be a weather index difference– again not tradable but only observable–, or it could be the GDP of a country.

In addition, it works similarly in a multidimensional set-up, preserving the gaussian distribution that is closed under marginalisation and conditioning. Hence, the solutions found under the Kullback-Leibler uncertainty maximization will be consistent across different dimensions.

When the underlying index is tradeable some identifications or comparisons can be made. Notice that the mean condition over the returns space $R_T$ will lead to the the well-known martingale condition $E_f[S_T p(0,T)] = S_0$ for asset pricing [Bjork, 2009], only when $\mu = \ln(1/p(0,T)) - \frac{\sigma^2}{2}$. Then the Kullback-Leibler uncertainty maximised distribution will coincide with the risk-neutral distribution derived under a complete market defined by the Black-Scholes model. In that case, $\sigma^2 = \sigma_{BS}^2(T-t)$.

**3.2. Multivariate set-up.** Furthermore, if now $\mathbf{R}_T \in \mathbb{R}^d$, that is we work in a multivariate set-up, with $d$ underlying asset having the vector of returns $\mathbf{R}_T = (R_T^{(1)}, \ldots, R_T^{(d)})$ then our problem to recover the distribution $f(\mathbf{R}_T)$ that minimizes the Kullback-Leibler divergence measure

$$\min I(f;g) = \min \int_{\mathbb{R}^d} f(\mathbf{R}_T) \ln \frac{f(\mathbf{R}_T)}{g(\mathbf{R}_T)} d\mathbf{R}_T \tag{21}$$

subject to the conditions

$$\int_{\mathbb{R}^d} f(\mathbf{R}_T) d\mathbf{R}_T = 1 \tag{22}$$

$$f(\mathbf{R}_T) > 0 \quad \forall \mathbf{R}_T \tag{23}$$

$$E_f(\mathbf{R}_T) = \mu \tag{24}$$

$$COVAR_f(\mathbf{R}_T) = \Sigma \tag{25}$$

again has a solution given by the multivariate gaussian distribution

$$f(\mathbf{R}_d) \propto \frac{1}{\sqrt{\det(\Sigma)}} \exp\left[-\frac{1}{2}(\mathbf{R}_T - \mu)^\top \Sigma^{-1}(\mathbf{R}_T - \mu)\right] \tag{26}$$

as proved by [Kapur, 1989].

**3.3. Option Pricing by Minimising the informational energy.** An informational energy concept was introduced by [Onicescu, 1966], see also [Perez, 1966]. The information energy associated with a probability density $p$ is defined by the functional

$$IE(p) = \int_{-\infty}^{\infty} p^2(x)dx \tag{27}$$

It is known that $IE$ is a convex functional and it is invariant under measure preserving transformations. Overall it has similar properties to the Shannon entropy and it has been introduced as a measure of randomness of a probability system. A discussion based on an axiomatic approach is discussed in [Theodorescu, 1977].

For the purposes of applications to finance one can assume without reduction of generality that the probability density behind the model is continuous. Furthermore, if one is also prepared to accept that the tails of the distribution given by $p$ decay to zero asymptotically one can prove, see [Calin and Udriste, 2014], the following result:

**Proposition 3.2.** *If $p$ is a probability density function on $\mathbb{R}$ such that it satisfied both conditions:*
(1) *$p$ is continuous*
(2) $\lim_{x\to\pm\infty} p(x) = 0$
*then $IE(p) = \int_{-\infty}^{\infty} p^2(x)dx < \infty$.*

[Gulko, 1999b] suggested that one may aim to work with the information energy functional defined in (27), although he has not called it on its name, as an alternative to entropy optimisation. In this case the problem we need to solve is

$$\min_f \int_{-\infty}^{\infty} f^2(R_T)dR_T \tag{28}$$

subject to

$$\begin{aligned} \int_{-\infty}^{\infty} f(R_T)dR_T &= 1 && (29)\\ f(R_T) &> 0 \quad \forall R_T && (30)\\ E_f(R_T) &= \mu && (31)\\ E_f(R_T^2) &= \sigma^2 - \mu^2 && (32) \end{aligned}$$

One can proceed again by Lagrangian optimisation.

$$\begin{aligned} \mathcal{L}(f(R_T), \lambda_1, \lambda_2) = IE(f) &+ \lambda_1\left(\mu - \int_{-\infty}^{\infty} R_T f(R_T)dR_T\right)\\ &+ \lambda_2\left(\sigma^2 - \mu^2 - \int_{-\infty}^{\infty} R_T^2 f(R_T)dR_T\right)\\ &+ \lambda_0\left(\int_{-\infty}^{\infty} f(R_T)dR_T - 1\right) + \int_{-\infty}^{\infty} f(R_T)\psi(R_T)dR_T \end{aligned}$$

that has the Lagrange multipliers $\lambda_1, \lambda_2, \lambda_0, \psi$. The dual problem is

$$\max_{\lambda,\lambda_0,\psi} \min_{f} \mathcal{L}(f(R_T), \lambda_1, \lambda_2) \tag{33}$$

which is a concave problem in parameters $\lambda_1, \lambda_2$ and $\psi$.

We need to solve out the Euler-Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial f} + \frac{d}{dR_T}\left(\frac{\partial \mathcal{L}}{\partial f'}\right) = 0 \tag{34}$$

The operator $\mathcal{L}(f, R_T)$ is independent of the derivative $f' = \frac{df(R_T)}{dR_T}$, the Euler-Lagrange equation is

$$2f + 1 - \lambda_1 R_T - \lambda_2 R_T^2 + \lambda_0 + \psi(R_T) = 0 \tag{35}$$

From the Karush-Kuhn-Tucker conditions we can impose $\psi(R_T) = 0$. Then the probability density function, up to a normalising constant, is given by

$$f(R_T) \propto \frac{1}{2}\left[\lambda_1 R_T + \lambda_2 R_T^2 - 1\right] \tag{36}$$

The problem here is that conditions (29) and (30) would be impossible to be satisfied given that the support of the sought probability distributions is $\mathbb{R}$. The solution comes from a more accurate interpretation, from a theoretical point of view of the underlying random variable $R_T$. Since this variable represents the percentage return of the underlying stock price $S_T$, the minimum value of this variable is clearly -1. The maximum value is mathematically infinite but financially it is also upper bounded, albeit the exact value of the upper bound $B$ may not be known. Therefore, we should search for distributions having support in an interval $[-1, B]$, where $B > 0$ is a very large number. Then

$$f(R_T) \propto \frac{\frac{1}{2}\left[\lambda_1 R_T + \lambda_2 R_T^2 - \lambda_0\right]}{\int_{-1}^{B} \frac{1}{2}\left[\lambda_1 x + \lambda_2 x^2 - \lambda_0\right] dx} \tag{37}$$

This is achieved by considering the Lagrangian operator

$$\begin{aligned} f(x) \quad \rightarrow \quad & \int_{-1}^{B} \mathcal{L}(f(x), x) dx \\ = \quad & \int_{-1}^{B} f^2(x) dx - \lambda_0 \left[\int_{-1}^{B} f(x) dx - 1\right] \\ & -\lambda_1 \left[\int_{-1}^{B} x f(x) dx - \mu\right] \\ & -\lambda_2 \left[\int_{-1}^{B} x^2 f(x) dx - (\sigma^2 - \mu^2)\right] \end{aligned}$$

Denoting by $\delta^2 \equiv \sigma^2 - \mu^2$, in short notation the Lagrangian operator is given by

$$\mathcal{L}(f, x) = f^2 - \lambda_0(f - 1) - \lambda_1(xf - \mu) - \lambda(x^2 f - \delta^2)$$

Hence, requiring that $\frac{\partial \mathcal{L}}{\partial f} = 0$ leads to the equation

$$2f - \lambda_0 - \lambda_1 x - \lambda_2 x^2 = 0$$

from which one obtains the solution (37).

The $\lambda$ parameters are obtained from the moment conditions. Thus, the parameters $\lambda$ are the solution to the following system of linear equations

$$\begin{cases} \lambda_0(B+1) + \frac{\lambda_1}{2}(B^2-1) + \frac{\lambda_2}{3}(B^3+1) = 2 \\ \lambda_0\frac{B^2-1}{2} + \frac{\lambda_1}{3}(B^3+1) + \frac{\lambda_2}{4}(B^4-1) = 2\mu \\ \lambda_0\frac{B^3+1}{2} + \frac{\lambda_1}{4}(B^4-1) + \frac{\lambda_2}{5}(B^5+1) = 2\delta^2 \end{cases}$$

This solution will have a unique solution if and only if the following determinant is different from zero

$$\begin{vmatrix} B+1 & \frac{1}{2}(B^2-1) & \frac{1}{3}(B^3+1) \\ \frac{1}{2}(B^2-1) & \frac{1}{3}(B^3+1) & \frac{1}{4}(B^4-1) \\ \frac{1}{3}(B^3+1) & \frac{1}{4}(B^4-1) & \frac{1}{5}(B^5+1) \end{vmatrix} \tag{38}$$

Since the coefficient of the highest order $B^9$ of this determinant is 0.0004629 we can conclude that for large enough $B$ the above determinant is positive. However, how large $B$ should be depends on a case by case basis.

## 4. Conclusions and Further Research

One of the main challenges linked to the entropy methods describing maximization of uncertainty is to work with a dynamic stochastic process rather than a one-period probability distribution. One possibility would be to maximize the uncertainty along a path, obtain a price, and then average all prices similarly to a Monte Carlo approach.

The multi-dimensional case also needs to be revisited. An equivalent condition to (20) should be derived for the multi-dimensional case.

Onicescu's information energy is also used to define a dependency concept that considers the similarity between the probability distributions of two random variables. This concept has not been used in financial mathematics so far and it should be researched in the future.

## References

[1] M. Avellaneda, C. Friedman, R. Holmes, D. Samperi, Calibrating volatility surfaces via relative-entropy minimization, *Applied Mathematical Finance* (1997) **4**, 37–64.

[2] M. Avellaneda, Minimum-entropy calibration of asset-pricing models, *International Journal of Theoretical and Applied Finance* (1998) **1**, 447–472.

[3] A.F. Bariviera , M.B. Guercio, L.B. Martinez, O.A. Rosso, Libor at crossroads: Stochastic switching detection using information theory quantifiers, *Chaos, Solitons and Fractals* (2016) *forthcoming*, 1–11.

[4] J. Borwein, R. Choksi, P. Marechal, Probability distributions of assets inferred from option prices via the principle of maximum entropy, *SIAM Journal on Optimization* (2003) **14**(2), 464–478.

[5] P.W. Buchen, M. Kelly, The maximum entropy distribution of an asset inferred from option prices, *Journal of Financial and Quantitative Analysis* (1996) **31**(1), 143–159.

[6] T. Bjork, *Arbitrage Theory in Continuous Time*, third edition, Oxford Finance Series, Oxford University Press, 2009.

[7] P.K. Bhatia, On Measures of Information Energy, *Information Sciences* (1997) **97**, 233–240.

[8] O. Călin, C. Udrişte, *Geometric Modeling in Probability and Statistics*, Springer, Berlin, 2014.

[9] I. Csiszar, Why Least Squares and Maximum Entropy? An Axiomatic Approach to Inference for Linear Inverse Problems, *Annals of Statistics* (1991) **19**, 2031–2066.

[10] T.M. Cover, J.A. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.

[11] C. Friedman, Y. Zhang, J. Huang, Estimating Univariate Distributions via Relative Entropy Minimization: Case Studies on Financial and Economic Data, *International Journal of Theoretical and Applied Finance* (2010) **13**(1), 163–193.
[12] N. Georgescu-Roegen, *The Entropy Law and the Economic Process*, Harvard University Press, Cambridge MA, 1971.
[13] A. Golan, G. Judge, D. Miller, *Maximum Entropy Econometrics: Robust Estimation with Limited Data*, John Wiley & Sons, New York, 1996.
[14] Les Gulko, The Entropy Theory of Bond Option Pricing, *International Journal of Theoretical and Applied Finance* (2002) **5**(4), 355–383.
[15] Les Gulko, The Entropy Theory of Stock Option Pricing, *International Journal of Theoretical and Applied Finance* (1999) **2**(3), 331–355.
[16] Les Gulko, The Entropic Market Hypothesis, *International Journal of Theoretical and Applied Finance* (1999) **2**(3), 293–329.
[17] G. Imbens, R. Spady, P. Johnson, Information Theoretic Approaches to Inference in Moment Condition Models, *Econometrica* (1998) **66**, 333–357.
[18] G.G. Judge, R.C. Mittelhammer, *An Information Theoretic Approach to Econometrics*, Cambridge University Press, Cambridge, 2012.
[19] J.N. Kapur, *Maximum Entropy Models In Science and Engineering*, Wiley, New York, 1989.
[20] S. Kullback, R.A. Leibler, On information and sufficiency, *Annals of Mathematical Statistics* (1951) **22**(1), 79–86.
[21] S. Kullback, *Information Theory and Statistics*, John Wiley & Sons, Chichester, 1959.
[22] D.G. Luenberger, *Optimization by Vector spaces Methods*, Wiley, New York, 1969.
[23] O. Onicescu, Energie informationnelle, *Comptes Rendus Hebdomadaires des Seances de l'Academie des Sciences Serie A* (1966) **263**, 841–842.
[24] L. Pardo, I.J. Taneja, Information energy and its Applications, In: (P. Hawkes ed.) *Advances in Electronic and Electron Physics*, Academic Press, 1991, 165–241.
[25] A. Perez, Sur l'energie informationnelle de M. Octavio Onicescu, *Revue Roumaine de Mathématique Pures et Appliquées* (1966) **72**, 1341–1347.
[26] A. Theodorescu, Energie Informationnelle et notions Apparentes, *Trabajos de Estadística e Investigación Operativa* (1977) **35**, 183–206.
[27] C.E. Shannon, A Mathematical Theory of Communication, *Bell System Technical Journal* (1948) **27**, 379–423.
[28] M. Stutzer, A simple nonparametric approach to derivative security valuation, *Journal of Finance* (1996) **51**(5), 1633–1652.
[29] H. Theil, *Economics and Information Theory*, North-Holland, Amsterdam, 1967.
[30] P. Theodossiou, Financial data and the skewed generalized t distribution, *Management Science* (1998) **44**(12), 1650–1661.
[31] R. Tunaru, Information Divergence: An Axiomatic Derivation, *Annals of the University of Craiova, Mathematics and Computer Science Series* (1991-1992) **19**, 51–54.

(Radu Tunaru) ICEQUFIN, University of Kent, Park Wood, Canterbury
*E-mail address*: R.Tunaru@kent.ac.uk