International Symposium on Information Theory and its Applications, ISITA2004
Parma, Italy, October 10–13, 2004

Two General Games Of Information

Flemming Topsøe

Department of Mathematics
University of Copenhagen,
Universitetsparken 5, 2100 COPENHAGEN, DENMARK
E-mail: topsoe@math.ku.dk

Abstract

The *Maximum Entropy Principle* (MaxEnt) as well as the *Minimum Information Divergence Principle* (MinDiv) and other optimization principles of information theory and its applications to physics, statistics, economy and other fields are here discussed from the standpoint of two-person zero-sum games.

1. INTRODUCTION

The recent publications [1] and [2] contain references to the previous development and do much to motivate the game theoretical approach which we shall adopt here.

Our model will be more general than found in previous work in the sense that it allows applications which go beyond standard notions of entropy and divergence (due to Shannon and to Kullback and Leibler). However, we simplify in another direction by imposing strong finiteness conditions so as to avoid infinite or undefined quantities in certain situations. Or rather, we argue *as if* such conditions are fulfilled, leaving it to the reader to relax the conditions when the need arises – as it often does, e.g. regarding applications to statistics.

Formally, we shall present two models, *absolute* and *relative* games. This will make the intended applications more clear but really, the relative games of Section 3 can be conceived as special cases of the absolute games of Section 2. Even more general models could be considered, as is clear from the indications in [3] (which has not yet been followed up by a full publication).

2. ABSOLUTE GAMES

We consider a quadruple $(\mathcal{S}_I, \mathcal{S}_{II}, i, \Phi)$ where $\mathcal{S}_I$ and $\mathcal{S}_{II}$ are sets, respectively the *potential strategy set for Player I* and the *strategy set for Player II*, $i : \mathcal{S}_I \to \mathcal{S}_{II}$ is a map of $\mathcal{S}_I$ into $\mathcal{S}_{II}$ and $\Phi : \mathcal{S}_I \times \mathcal{S}_{II} \to \mathbb{R}$, denoted $(P, Q) \curvearrowright \Phi(P\|Q)$, is a real function, the *complexity function*[1]. For the most natural applications, the complexity function will be non-negative. However, in order to enable a reduction of the study of the relative games of the next section to the case of absolute games introduced here, we allow general real-valued complexity functions.

If $(P, Q) \in \mathcal{S}_I \times \mathcal{S}_{II}$ and $Q = i(P)$, we also write $Q = \hat{P}$. As i is often understood and in most applications even very simple, viz. the identity map (thus, in such cases, $\mathcal{S}_I = \mathcal{S}_{II}$), we may focus on the triple $(\mathcal{S}_I, \mathcal{S}_{II}, \Phi)$ or just on the function Φ.

Our basic assumptions are that

$$\Phi(P\|Q) \geq \Phi(P\|\hat{P}) \text{ for all } (P, Q) \in \mathcal{S}_I \times \mathcal{S}_{II}$$

and that here, equality holds if and only if $Q = \hat{P}$.

We define Φ-*entropy*, $H_\Phi : \mathcal{S}_I \to \mathbb{R}$, and Φ-*divergence*, $D_\Phi : \mathcal{S}_I \times \mathcal{S}_{II} \to \mathbb{R}$, by

$$H_\Phi(P) = \inf_{Q \in \mathcal{S}_{II}} \Phi(P\|Q),$$

$$D_\Phi(P\|Q) = \Phi(P\|Q) - H_\Phi(P).$$

By assumption,

$$H_\Phi(P) = \Phi(P\|\hat{P}) \text{ for all } P \in \mathcal{S}_I .$$

Now, let $\mathcal{P}$, the *actual strategy set for Player I*, be a given subset of $\mathcal{S}_I$ and consider the two-person zero-sum game $\gamma_\Phi(\mathcal{P})$ with $\mathcal{P}$ and $\mathcal{S}_{II}$ as strategy sets for the two players and with Φ as objective function, conceived as a cost to Player II. Then, for Player I, a strategy $P \in \mathcal{P}$ is *optimal* if $H_\Phi(P) = H_\Phi^{\max}(\mathcal{P})$, the *MaxEnt-value* which is defined by

$$H_\Phi^{\max}(\mathcal{P}) = \sup_{P \in \mathcal{P}} \mathrm{H}_\Phi(P).$$

This work was supported by INTAS, project 00-738 and by the Danish Natural Science Research Council.

[1] Some may prefer to refer to this function as *description length* "à la Rissanen", cf. [4]. Anyhow, the interpretation is that $\Phi(P\|Q)$ measures the difficulty ("complexity") involved when the "system" is in "state" P chosen by "nature" (Player I!) and when "we" (Player II!) use Q as the tool for "observation" or "description" of the state of the system.

And, for Player II, a strategy $Q \in \mathcal{S}_{II}$ is *optimal* if the *risk* associated with $Q \in \mathcal{S}_{II}$, $\mathrm{R}_\Phi(Q|\mathcal{P})$, equals the *minimal risk*, $\mathrm{R}_\Phi^{\min}(\mathcal{P})$, with these quantities defined by

$$\mathrm{R}_\Phi(Q|\mathcal{P}) = \sup_{P\in\mathcal{P}} \Phi(P\|Q),$$

$$\mathrm{R}_\Phi^{\min}(\mathcal{P}) = \inf_{Q\in\mathcal{S}_{II}} \mathrm{R}_\Phi(Q|\mathcal{P}).$$

If misunderstanding is unlikely, we may write $\mathrm{R}_\Phi(Q)$ instead of $\mathrm{R}_\Phi(Q|\mathcal{P})$.

Clearly,

$$H_\Phi^{\max}(\mathcal{P}) \le \mathrm{R}_\Phi^{\min}(\mathcal{P}),$$

and, by definition, the game $\gamma_\Phi(\mathcal{P})$ is in *equilibrium* if equality holds here. We find it convenient also to introduce the *excess* which we, ignoring dependence on Φ and $\mathcal{P}$, simply denote by ε. It is given by

$$\varepsilon = \mathrm{R}_\Phi^{\min}(\mathcal{P}) - H_\Phi^{\max}(\mathcal{P})\,.$$

Lemma 1. *Consider the game $\gamma_\Phi(\mathcal{P})$ and let ε be the excess.*

(i). If Player I has an optimal strategy P^*, then, for every Player II-strategy Q,

$$\mathrm{R}_\Phi^{\min}(\mathcal{P}) + \mathrm{D}(P^*\|Q) \le \varepsilon + \mathrm{R}_\Phi(Q)\,. \tag{1}$$

(ii). If Player II has an optimal strategy Q^*, then, for every Player I-strategy P,

$$H(P) + \mathrm{D}(P\|Q^*) \le \varepsilon + H_\Phi^{\max}(\mathcal{P})\,. \tag{2}$$

Proof. (i): Assume that $\mathrm{H}_\Phi(P^*) = H_\Phi^{\max}(\mathcal{P})$. Then, for any $Q \in \mathcal{S}_{II}$,

$$\begin{aligned}\mathrm{R}_\Phi^{\min}(\mathcal{P}) + \mathrm{D}(P^*\|Q) &= \varepsilon + H_\Phi^{\max}(\mathcal{P}) + \mathrm{D}(P^*\|Q)\\ &= \varepsilon + \mathrm{H}_\Phi(P^*) + \mathrm{D}(P^*\|Q)\\ &= \varepsilon + \Phi(P^*\|Q)\\ &\le \varepsilon + \mathrm{R}_\Phi(Q)\,,\end{aligned}$$

and (i) follows.

(ii): Assume that $\mathrm{R}_\Phi(Q^*) = \mathrm{R}_\Phi^{\min}(\mathcal{P})$. Then, for any $P \in \mathcal{S}_I$,

$$\begin{aligned}H(P) + \mathrm{D}(P\|Q^*) &= \Phi(P\|Q^*) \le \mathrm{R}_\Phi(Q^*)\\ &= \mathrm{R}_\Phi^{\min}(\mathcal{P}) = \varepsilon + H_\Phi^{\max}(\mathcal{P})\,,\end{aligned}$$

hence (ii) follows. □

For the further study of the game $\gamma_\Phi(\mathcal{P})$, the notion of a *Nash equilibrium pair* for $\gamma_\Phi(\mathcal{P})$ is important. This is a pair $(P^*, Q^*) \in \mathcal{P} \times \mathcal{S}_{II}$ such that, for every $(P, Q) \in \mathcal{P} \times \mathcal{S}_{II}$, the *saddle value inequalities*

$$\Phi(P\|Q^*) \le \Phi(P^*\|Q^*) \le \Phi(P^*\|Q)$$

hold. When these inequalities hold, no player can benefit from changing his strategy (P^* or Q^*) provided the other player stick to his (Q^* or P^*). The significance in our setting is made clear in Theorem 1 of Section 4.

3. RELATIVE GAMES

Consider a model $(\mathcal{S}_I, \mathcal{S}_{II}, i, Q_0, \mathrm{D})$ with $(\mathcal{S}_I, \mathcal{S}_{II}, i)$ as in Section 2, with Q_0, the *reference* (or *prior*) an element in $\mathcal{S}_{II}$ and with $\mathrm{D} : \mathcal{S}_I \times \mathcal{S}_{II} \to \mathbb{R}_+$, the *divergence function*, a function such that

$$\mathrm{D}(P\|Q) \ge 0$$

for $(P, Q) \in \mathcal{S}_I \times \mathcal{S}_{II}$ and such that equality holds if and only if $Q = \hat{P}$.

Let $(P, Q) \in \mathcal{S}_I \times \mathcal{S}_{II}$ be a pair of possible strategies for the two players. We then define the associated *calibration gain*, denoted $\mathrm{D}(P\|Q_0 \rightsquigarrow Q)$, as the quantity

$$\mathrm{D}(P\|Q_0 \rightsquigarrow Q) = \mathrm{D}(P\|Q_0) - \mathrm{D}(P\|Q).$$

For $\mathcal{P} \subseteq \mathcal{S}_I$, the *relative game* $\gamma_\mathrm{D}(\mathcal{P}, Q_0)$ is defined as the two-person zero-sum game with $\mathcal{P}$ and $\mathcal{S}_{II}$ as strategy sets for the two players and with calibration gain restricted to $\mathcal{P} \times \mathcal{S}_{II}$ as objective function, this time conceived as a gain from the point of view of Player II.

Here, key quantities to consider are the *MinDiv-value* (minimum divergence) given by

$$\begin{aligned}\mathrm{D}^{\min}(\mathcal{P}|Q_0) &= \inf_{P\in\mathcal{P}} \sup_{Q\in\mathcal{S}_{II}} \mathrm{D}(P\|Q_0 \rightsquigarrow Q)\\ &= \inf_{P\in\mathcal{P}} \mathrm{D}(P\|Q_0)\end{aligned}$$

and the *MaxGain-value* (maximal gain). The latter is defined as the supremum of the *guaranteed gains* associated with fixed strategies. Using the letter "Γ"for "gain" , the basic definitions are as follows:

$$\Gamma_\mathrm{D}^{\max}(\mathcal{P}|Q_0) = \sup_{Q\in\mathcal{S}_{II}} \Gamma_\mathrm{D}(Q|\mathcal{P}, Q_0)$$

with

$$\Gamma_\mathrm{D}(Q|\mathcal{P}, Q_0) = \inf_{P\in\mathcal{P}} \mathrm{D}(P\|Q_0 \rightsquigarrow Q).$$

Now,

$$\Gamma_\mathrm{D}^{\max}(\mathcal{P}|Q_0) \le \mathrm{D}^{\min}(\mathcal{P}|Q_0).$$

If equality holds above, the game $\gamma_\mathrm{D}(\mathcal{P}, Q_0)$ is in *equilibrium.*

A strategy $P \in \mathcal{P}$ is *optimal* for Player I if $D(P\|Q_0) = \mathrm{D}^{\min}(\mathcal{P}|Q_0)$ and a strategy $Q \in \mathcal{S}_{II}$ is *optimal* for Player II if $\Gamma_\mathrm{D}(Q|\mathcal{P}, Q_0) = \Gamma_\mathrm{D}^{\max}(\mathcal{P}|Q_0)$.

A pair $(P^*, Q^*) \in \mathcal{P} \times \mathcal{S}_{II}$ is a *Nash equilibrium pair* for $\gamma_\mathrm{D}(\mathcal{P}, Q_0)$ if the inequalities

$$D(P^*\|Q_0 \rightsquigarrow Q) \le \mathrm{D}(P^*\|Q_0 \rightsquigarrow Q^*) \le \mathrm{D}(P\|Q_0 \rightsquigarrow Q^*)$$

hold for all $(P, Q) \in \mathcal{P} \times \mathcal{S}_{II}$.

4. EQUILIBRIUM AND OPTIMAL STRATEGIES

The following general result holds:

Theorem 1. *(i). The absolute game $\gamma_\Phi(\mathcal{P})$ is in equilibrium and both players have optimal strategies if and only if there exists a Nash equilibrium pair (P^*, Q^*) for the game, and this holds if and only if there exists $P^* \in \mathcal{P}$ such that*

$$\Phi(P\|\hat{P}^*) \le \mathrm{H}_\Phi(P^*) \text{ for all } P \in \mathcal{P}\,.$$

If so, then $Q^* = \hat{P}^*$ and for any pair $(P, Q) \in \mathcal{P} \times \mathcal{S}_{II}$, the following inequalities hold:

$$\mathrm{H}_\Phi(P) + \mathrm{D}_\Phi(P\|Q^*) \le H_\Phi^{\max}(\mathcal{P}) \tag{3}$$

$$\le \mathrm{R}_\Phi^{\min}(\mathcal{P}) \le \mathrm{R}_\Phi(Q) - \mathrm{D}_\Phi(P^*\|Q)\,. \tag{4}$$

(ii). The relative game $\gamma_\mathrm{D}(\mathcal{P}, Q_0)$ is in equilibrium and both players have optimal strategies if and only if there exists a Nash equilibrium pair (P^*, Q^*) for the game, and this holds if and only if there exists $P^* \in \mathcal{P}$ such that

$$\mathrm{D}(P^*\|Q_0) \le \mathrm{D}(P\|Q_0 \rightsquigarrow \hat{P}^*) \text{ for all } P \in \mathcal{P}\,.$$

If so, then $Q^* = \hat{P}^*$ and for $(P, Q) \in \mathcal{P} \times \mathcal{S}_{II}$, the following inequalities hold:

$$\mathrm{D}(P\|Q_0) - \mathrm{D}(P\|\hat{P}^*) \ge \mathrm{D}^{\min}(\mathcal{P}|Q_0) \tag{5}$$

$$\ge \mathrm{R}_\mathrm{D}^{\max}(\mathcal{P}|Q_0) \ge \Gamma_\mathrm{D}(Q|\mathcal{P}, Q_0) + \mathrm{D}(P^*\|Q)\,. \tag{6}$$

Proof. (i): The simple proof is given in [1]. Basic parts may be derived from Lemma 1.

(ii): Consider Φ given by

$$\Phi(P\|Q) = -\,\mathrm{D}(P\|Q_0 \rightsquigarrow Q)\,,$$

and note that $\mathrm{H}_\Phi(P) = -\,\mathrm{D}(P\|Q_0)$, $\mathrm{D}_\Phi(P\|Q) = \mathrm{D}(P\|Q_0)$ and $\mathrm{R}_\Phi(Q) = -\Gamma_\mathrm{D}(Q|\mathcal{P}, Q_0)$. Then apply (i) to Φ and the results of (ii) follow. □

A pair (P^*, Q^*) is *the bi-optimal matching pair* for the game considered (either $\gamma_\Phi(\mathcal{P})$ or $\gamma_\mathrm{D}(\mathcal{P}, Q_0)$) if P^* is an optimal strategy for Player I, Q^* is an optimal strategy for Player II and $Q^* = \hat{P}^*$ and (P^*, Q^*) is the unique pair with these properties. Theorem 1 tells us that if a Nash equilibrium pair exists, then so does the bi-optimal matching pair and, furthermore, the game in question is in equilibrium in the sense defined previously. The existence of the bi-optimal matching pair is a relatively weak notion of equilibrium which it appears worth while to study further. If i is the identity map, we say that P^* is *the bi-optimal strategy* if (P^*, P^*) is the bi-optimal matching pair.

Further theoretical investigations depend on the introduction of more structure on the strategy sets $\mathcal{S}_I$ and $\mathcal{S}_{II}$. Typically, this involves topological and linear or affine structure and convexity considerations become important. Though the optimal abstract setting for this kind of modelling is perhaps not yet in place, a good idea about the possibilities may be gathered from [1], [2], [3] and references therein.

Modelling in a probabilistic context is the most common. However, wider possibilities exist, not necessarily related to information theory. As an indication, see Subsections 5.1 and 5.2 below.

5. EXAMPLES

Apart from the first two examples, $\mathcal{S}_I$ and $\mathcal{S}_{II}$ are sets of probability distributions or sets closely related to such spaces. We ignore below that strict finiteness conditions as required in Sections 2 and 3 may not always hold.

5.1. A Problem of Location

A classical problem of location theory, cf. [5], was introduced as follows by Sylvester in 1857: "It is required to find the least circle which shall contain a given system of points in a plane". If we take $\mathcal{S}_I = \mathcal{S}_{II} = \mathbb{R}^2$ and $\Phi(P\|Q) = \|P - Q\|$, the Euclidean distance between P and Q, and consider a suitable set $\mathcal{P}$ of points, then Sylvesters problem corresponds to that of finding an optimal strategy for Player II in the associated absolute game. Clearly, here every strategy for Player I is optimal and the game is not in equilibrium. Anyhow, if $\mathcal{P}$ is convex with finitely many extremal points, the bi-optimal distribution exists. If we change the strategy set for Player I by allowing randomization, the situation changes and Theorem 1 can be applied.

5.2. Convex Sets and Nearest Points

If, for the problem above, one considers relative games, one will more often find games in equilibrium. One will observe that whereas a monotone transformation of the distance function will not influence the question of equilibrium for the absolute game, such a change may well effect the relative games in a significant way. It is appropriate to choose squared norms. For instance, one may take $\mathcal{S}_I = \mathcal{S}_{II}$ to be a Hilbert space, and D to be defined by $\mathrm{D}(P\|Q) = \|P - Q\|^2$. For closed convex sets $\mathcal{P}$ and any reference point Q_0, equilibrium holds and the orthogonal projection of Q_0 on $\mathcal{P}$ is the bi-optimal strategy.

This example and the previous one indicate that interesting games which are not in equilibrium do occur, that squared metrics may be appropriate to consider and that it is more natural for Player II to have unique optimal strategies than for Player I.

5.3. The MaxEnt Principle

Let $\mathbb{A}$ be finite, put $\mathcal{S}_I = M_+^1(\mathbb{A})$, the set of probability distributions over $\mathbb{A}$, and put $\mathcal{S}_{II} = K(\mathbb{A})$, the set of (abstract) *codes* over $\mathbb{A}$ defined as the set of $\kappa : \mathbb{A} \to [0, \infty]$ such that *Kraft's equality*

$$\sum_{a \in \mathbb{A}} e^{-\kappa(a)} = 1$$

holds.

The map i is given by $\kappa(a) = -\ln P(a)\,;\, a \in \mathbb{A}$ where $\kappa = i(P) = \hat{P}$.

If we take Φ as *averagge code length*:

$$\Phi(P\|\kappa) = \langle \kappa, P \rangle = \sum_{a \in \mathbb{A}} P(a)\kappa(a)\,,$$

and consider the associated absolute game, we are led to the MaxEnt-principle. See [6] for a rather comprehensive study from the game theoretical point of view.

5.4. The MinDiv Principle

Now take $\mathcal{S}_I = \mathcal{S}_{II} = M_+^1(\mathbb{A})$ and as D choose *Kullback-Leibler divergence*

$$\mathrm{D}(P\|Q) = \sum_{a \in \mathbb{A}} P(a) \ln \frac{P(a)}{Q(a)}\,.$$

When we fix some prior distribution as our reference, we may consider the associated relative game. This leads us to the MinDiv principle and the much studied notion of *information projection*, cf. [7] and further references in [2]. Also look out for the soon-to-appear comprehensive treatment [8].

The inequalities (3)-(6) of Theorem 1 are *Pythagorean type inequalities*. The game theoretical derivation of these inequalities appears illuminating, e.g. the derivation indicates that the most well known of the inequalities, viz. (5), which is mainly associated with Csiszár's name, involves strategies for Player I but is really more closely associted with the existence of an optimal strategy for Player II, cf. also Lemma 1.

5.5. Channel Capacity

Consider a discrete memoryless channel defined in terms of a finite *input alphabet* $\mathbb{A}$, a finite *output alphabet* $\mathbb{B}$ and a *Markov kernel* $\mathbb{P}$. Take $\mathcal{S}_I$ to be the set of input distributions, $\mathcal{S}_{II}$ to be the set of output distributions and let i map an input distribution to the induced output distribution. As Φ take *information transmission rate*. The absolute game then leads to the *capacity-redundancy* theorem. More details are in [2].

5.6. Non-extensive Entropy

The group of examples discussed here is related to areas of statistical physics where there is a need to go beyond Shannon entropy. For details, see [1].

Consider $\mathcal{S}_I = \mathcal{S}_{II} = M_+^1(\mathbb{A})$ with $\mathbb{A}$ finite, say, and look at complexity functions which induce *Csiszár f-divergences*. We shall assume that f is smooth, say twice differentiable on $]0, \infty[$, that f is strictly convex on $[0, \infty[$ and that $f(0) = f(1) = 0$ and $f'(1) = 1$. Such functions are here called *generators*.

Using standard conventions and denoting point-probabilities of P and Q by (p_i), respectively (q_i), we define $\Phi = \Phi_f$ by

$$\Phi_f(P\|Q) = \sum_{i \in \mathbb{A}} \big(q_i f(\frac{p_i}{q_i}) - f(p_i)\big)\,.$$

Then, writing H_f for H_Φ and D_f for D_Φ,

$$\mathrm{H}_f(P) = -\sum_{i \in \mathbb{A}} f(p_i)\,,$$

$$\mathrm{D}_f(P\|Q) = \sum_{i \in \mathbb{A}} q_i f(\frac{p_i}{q_i})\,.$$

Expressed in terms of the *Csiszár-dual* defined by

$$\widetilde{f}(x) = x f(\frac{1}{x})\,,\;\; 0 \le x \le \infty\,,$$

we find that

$$\Phi_f(P\|Q) = \sum_{i \in \mathbb{A}} p_i \Big(\widetilde{f}(\frac{q_i}{p_i}) - \widetilde{f}(\frac{1}{p_i})\Big)\,,$$

$$\mathrm{H}_f(P) = -\sum_{i \in \mathbb{A}} p_i \widetilde{f}(\frac{1}{p_i})\,,$$

$$\mathrm{D}_f(P\|Q) = \sum_{i \in \mathbb{A}} p_i \widetilde{f}(\frac{q_i}{p_i})\,.$$

These formulas may lead to interesting interpretations which extend the standard interpretations for the classical case which corresponds to the choice $f(x) = x \ln x$, $\widetilde{f}(x) = \ln \frac{1}{x}$.

A concrete two-parameter family of generators is obtained by taking

$$f_{\alpha,\beta}(x) = x \ln_{\alpha,\beta}(x)$$

where the *deformed logarithms* occurring here are defined by

$$\ln_{\alpha,\beta} x = \frac{x^\beta - x^\alpha}{\beta - \alpha} \text{ if } \beta \neq \alpha$$

($x^\alpha \ln x$ if $\beta = \alpha$).

By far the most popular subfamily among physicists was introduced in 1988 by Tsallis, cf. [9], and corresponds to the family $(f_{q-1,0})_{q>0}$. This leads to a family of entropy measures closely related to Rényi entropy. For $q = \frac{1}{2}$, *Hellinger divergence*, a squared metric, appears (it is the only symmetric divergence measure in the Tsallis family).

In [1] the following result was proved:

Theorem 2. *If $0 < q \leq 1$ (and for no other parameter values), the absolute games associated with Φ_f ; $f = f_{q-1,0}$ are in equilibrium for any compact convex set $\mathcal{P} \subseteq M_+^1(\mathbb{A})$ and the bi-optimal distribution exists.*

5.7. Calibration in a Model from Finance

We turn to an application to mathematical finance, cf. [10], which involves the following standard model of stock price development: $(S_k)_{k\geq 0}$ where the price at time k is given by $S_k = S_{k-1}(1+\rho_k)$. Here, $S_0 > 0$ is deterministic and $(\rho_k)_{k\geq 1}$ an iid sequence with values in $]-1,\infty[$. A prior distribution, Q_0, of the ρ_k's is given. This may not render $(S_k)_{k\geq 0}$ a martingale w.r.t. the natural filtration of σ-fields. Therefore, the model based on Q_0 may not respect the *no arbitrage principle* and a move to a new state, P, can be expected.

The martingale condition amounts to the vanishing of the mean value of the ρ_k's in the new model based on P. By the assumptions, this does not depend on k. We may then translate everything to conditions involving only distributions on $\mathbb{R}$. We only consider discrete distributions with finite support. Let $\langle \cdot, P\rangle$ denote mean value w.r.t. P. We denote by *id* the identity map on $\mathbb{R}$.

For the prior Q_0 we assume, say, that $\langle id, Q_0\rangle < 0$, i.e. there is a *negative trend*, and that some point in the support of Q_0 is positive. The set $\mathcal{P}$ of acceptable market states is modelled by distributions P such that $\langle id, P\rangle = 0$ and, with an exception discussed later, we also demand that P is equivalent to Q_0. We have $\mathcal{S}_I \neq \emptyset$, i.e. the market is *incomplete*. The set $\mathcal{S}_{II}$, referred to as the set of *calibration strategies* is taken as the set of distributions which are absolutely continuous w.r.t. Q_0. As divergence function Φ we only consider one choice here, viz. *reverse relative entropy* given by

$$\mathrm{D}(P\|Q) = \sum Q(a) \ln \frac{Q(a)}{P(a)} .$$

A bit surprisingly, the following result holds:

Theorem 3. *The game $\gamma_{\mathrm{D}}(\mathcal{P}, Q_0)$ is in equilibrium and the bi-optimal calibration strategy exists if and only if the support of Q_0 contains only one positive element.*

It is possible to identify the bi-optimal distribution. This we shall do in the simple case when Q_0 is supported by $-a, 0, a$ where $a > 0$. Let the point probabilities of Q_0 be (p, q, r). Then the bi-optimal distribution is given by the point probabilities $\left(\frac{1}{2}(p+r), q, \frac{1}{2}(p+r)\right)$, thus involves arithmetic averages (whereas usual divergence would lead to geometric averages).

We point out that calibration using reverse relative entropy makes it possible to calibrate meaningfully when $r = 0$. The calibrated distribution has point probabilities $(\frac{p}{2}, q, \frac{p}{2})$. (Calibration w.r.t. relative entropy would give the unreasonable result $P^* = (0,1,0)$). Calibration when $r = 0$ corresponds to a situation with negative trend when previous evidence did not show any instance of increase in stock price but, never the less, for one reason or another, one expects that increase in stock price could occur in the future but only corresponding to the location a.

References

[1] F. Topsøe. Entropy and Equilibrium via Games of Complexity. *Physica A*, 340/1-3:11–31, 2004.

[2] F. Topsøe. Information Theory at the Service of Science. 2004. To appear in a special volume of the János Bolyai Mathematical Society, available at http://www.math.ku.dk/∼topsoe.

[3] P. Harremoës and F. Topsøe. Unified approach to optimization techniques in Shannon theory. In *Proceedings, 2002 IEEE International Symposium on Information Theory*, page 238. IEEE, 2002.

[4] P. D. Grünwald. A Tutorial Introduction to the Minimum Description Length Principle. In Myung Grünwald and Pitt, editors, *Advances in Minimum Description Length: Theory and Applications.* MIT Press, Bosten, 2004. to appear.

[5] Z. Drezner and H. Hamacher, editors. *Facility location. Applications and Theory.* Springer, Berlin, 2002.

[6] P. Harremoës and F. Topsøe. Maximum Entropy Fundamentals. *Entropy*, (3):191 – 226, Sept. 2001. ONLINE http://www.mdi.org/entropy/.

[7] I. Csiszár. I-divergence geometry of probability distributions and minimization problems. *Ann. Probab.*, 3:146–158, 1975.

[8] P. D. Grünwald and A. P. Dawid. Game Theory, Maximum Entropy, Minimum Discrepancy, and Robust Bayesian Decision Theory. 2004. to appear in Annals of Statistics.

[9] C. Tsallis. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Physics*, 52:479, 1988. See http://tsallis.cat.cbpf.br/biblio.htm for a complete and updated bibliography.

[10] Z. Khechinashvili O. Glonti, P. Harremoës and F. Topsøe. Nash equilibrium in a game of calibration. 2004. available at http://www.math.ku.dk/∼topsoe.