Interaction between Truth and Belief as the key to non-extensive statistical physics

Flemming Topsøe
University of Copenhagen, Department of Mathematical Sciences
Universitetsparken 5, 2100 Copenhagen, Denmark
Presentation at SigmaPhi2008

Claim : For the first time (?) the major part of a reasonable interpretation of the class of Tsallis entropies is provided which can explain their significance. This rests on the primary assumption that truth and belief interact and on a natural variational principle.

Our approach

– is philosophical: Put yourself in the shoes of the physicist who is planning observations and see if you can accept the considerations below
(numbered **1 – 8**).

1 Events have truth-assignments and belief-assignments, respectively x and y. These are numbers in $[0, 1]$, hence may be conceived as probabilities.

2 Any event I may observe entails a certain effort on my part. Before embarking on observations, I will determine the effort which I am willing to or have to devote to any event I may be faced with. This effort must only depend on my belief, y, and is denoted by $\kappa(y)$. The function κ, is the coder. As 1 represents certainty, $\kappa(1) = 0$.

3 To determine the coder, I must know the basic characteristics of the world I operate in. I will focus primarily on interaction between truth and belief.

4 I will model the interaction by a function π, the interaction, defined on $[0,1] \times [0,1]$. My idea is that $\pi(x, y)$ represents the weight with which the world will present an event to me in case the truth-assignment is x and my belief in the event is y.

4 I will model the interaction by a function π, the interaction, defined on $[0,1] \times [0,1]$. My idea is that $\pi(x,y)$ represents the weight with which the world will present an event to me in case the truth-assignment is x and my belief in the event is y.

Schematically, with $\pi_i = \pi(x_i, y_i)$:

$\mathbb{A}$	Truth	Belief	Interaction
.	.	.	.
.	.	.	.
.	.	.	.
i	x_i	y_i	π_i
.	.	.	.
.	.	.	.
.	.	.	.

Example: The classical world is a world of "no interaction", hence the interaction is $\pi(x,y) = x$.

5 I believe that my world is consistent in the sense that $\sum_{i\in\mathbb{A}} \pi_i = 1$ whenever $(x_i)_{i\in\mathbb{A}}$ and $(y_i)_{i\in\mathbb{A}}$ are probability assignments and $\pi_i = \pi(x_i, y_i)$.

Note: Then interaction must be sound,
i.e. $\pi(x, x) = x$ for all $x \in [0, 1]$.

6 To enable observations I must configure available observation- and measuring devices.

The resulting configuration will enable me to perform experiments, i.e. to study particular situations (physical systems) from the world which have my interest.

7 Separability applies: My total effort related to observations from the configured situation is the sum of individual contributions. Weights must be assigned to each contribution according to the weight with which I will experience the various events. The total effort I also refer to as the complexity , Φ. Thus:

$$\Phi(x, y) = \sum_{i \in \mathbb{A}} \pi(x_i, y_i)\kappa(y_i)$$

with $x = (x_i)_{i \in \mathbb{A}}$ the truth- and $y = (y_i)_{i \in \mathbb{A}}$ the belief-assignments associated with the events.

8 I will attempt to minimize complexity and shall appeal to the principle that complexity is the smallest when belief matches truth, ($(y_i)_{i\in\mathbb{A}} = (x_i)_{i\in\mathbb{A}}$). As

$$\sum_{i\in\mathbb{A}} \pi(x_i, y_i)\kappa(y_i) - \sum_{i\in\mathbb{A}} x_i\kappa(x_i)$$

represents my frustration, the principle says that frustration is the least, in fact disappears, when $(y_i)_{i\in\mathbb{A}} = (x_i)_{i\in\mathbb{A}}$.

Note: Given $x = (x_i)_{i\in\mathbb{A}}$, minimal complexity is what I am aiming at. It is an important quantity. I will call it entropy:

$$\mathsf{S}(x) = \inf_{y=(y_i)_{i\in\mathbb{A}}} \Phi(x, y) = \sum_{i\in\mathbb{A}} x_i\kappa(x_i)\,.$$

Frustration too looks important. Perhaps I better call it divergence:

$$\mathsf{D}(x, y) = \Phi(x, y) - \mathsf{S}(x)\,.$$

To summarize:

1: Events have Truth- and Belief- assignments (x and y).

2: Events emply effort on my part, $\kappa(y)$ (κ is the coder).

3: Characteristic of my world: Interaction btw. Truth and Belief.

4: Interaction $\pi(x, y)$ gives weight with which I will see a Truth-x, Belief-y event.

5: World is consistent: $\sum \pi(x_i, y_i) = 1$ when ...

6: I must configure devices to enable observation.

7: Total effort, complexity , is $\sum \pi(x_i, y_i)\kappa(y_i)$.

8: Frustration $\sum \pi(x_i, y_i)\kappa(y_i) - \sum x_i\kappa(x_i)$ disappears when Belief matches Truth (and not otherwise).

Can you accept all this? If so, you can conclude:

Theorem: Modulo regularity conditions and a condition of normalization, $q = \pi(1, 0)$ must be non-negative and π and κ uniquely determined from q by:

$$\pi(x, y) = qx + (1 - q)y \,, \tag{1}$$

$$\kappa(y) = \ln_q \frac{1}{y} \,, \tag{2}$$

where the q-logarithm is given by

$$\ln_q x = \begin{cases} \ln x \text{ if } q = 1, \\ \frac{x^{1-q}-1}{1-q} \text{ if } q \neq 1 \,. \end{cases}$$

Hence entropy is given by

$$S(x) = \sum_{i \in \mathbb{A}} x_i \ln_q \frac{1}{x_i} \,.$$

This is the essence of my contribution. Can *you*, physicists in particular, contribute to illuminate key outstanding issues (or point to already existing relevant results):

Challenges:

- explain interaction on physical grounds,
- suggest possibilities for an accompanying process of real coding,
- illuminate the good sense (if any :-)) of the views put forward in well studied concrete cases (possibly distinguishing between the cases $0 < q < 1$, $1 < q \leq 2$ and $q > 2$).

If time permits, let us look into the following:

- proof of theorem
- connection with Bregman generation
- relaxing the condition of consistency.

Indication of proof of main result

Functions π and κ are assumed continuous on their domains and continuously differentiable and finite valued on the interiors of their domains. Normalization of κ means that $\kappa(1) = 0$ and that $\kappa'(1) = -1$.

You can exploit the consistency condition to show that, for all $(x, y) \in [0, 1]^2$,

$$\pi(x, y) = qx + (1 - q)y$$

with $q = \pi(1, 0)$.

Consider a fixed finite probability vector $(x_i)_{i \in \mathbb{A}}$ with all x_i positive. Varying $(y_i)_{i \in \mathbb{A}}$ we find, via the introduction of a Lagrange multiplier, that f given by

$$f(x) = \frac{\partial \pi}{\partial y}(x, x)\kappa(x) + \pi(x, x)\kappa'(x)$$

is constant on $\{x_i | i \in \mathbb{A}\}$. Exploiting this for three-element alphabets $\mathbb{A}$ shows that $f \equiv -1$. Then the formula for κ is readily derived.

Bregman generation: Look at concave generator h_q and associated "Bregman quantities":

$$\begin{cases} h_q(x) = x \ln_q \frac{1}{x}\,, \\ \phi_q(x,y) = h_q(y) + (x-y)h_q'(y)\,, \\ d_q(x,y) = h_q(y) - h_q(x) + (x-y)h_q'(y)\,, \\ \Phi_q(P,Q) = \sum_{i\in\mathbb{A}} \phi_q(p_i,q_i)\,, \\ \mathsf{S}_q(P) = \sum_{i\in\mathbb{A}} h_q(p_i)\,, \\ \mathsf{D}_q(P,Q) = \sum_{i\in\mathbb{A}} d_q(p_i,q_i)\,. \end{cases}$$

-compare with "interaction quantities":

$$\begin{cases} \pi_q(x,y) = qx + (1-q)y \text{ (interaction)}\,, \\ \kappa_q(x) = \ln_q \frac{1}{x} \text{ (coder)}\,, \\ \xi(x,y) = y - x\,, \text{ (corrector)}\,, \\ \Phi_q(P,Q) = \sum_{i\in\mathbb{A}} \pi_q(p_i,q_i)\kappa_q(q_i) \\ \quad = \sum_{i\in\mathbb{A}} \Big(\pi_q(p_i,q_i)\kappa_q(q_i) + \xi(p_i,q_i)\Big)\,, \\ \mathsf{S}_q(P) = \sum_{i\in\mathbb{A}} p_i\kappa_q(p_i)\,, \\ \mathsf{D}_q(P,Q) = \sum_{i\in\mathbb{A}} \Big(\pi_q(p_i,q_i)\kappa_q(q_i) - p_i\kappa_q(p_i)\Big) \\ \quad = \sum_{i\in\mathbb{A}} \Big(\pi_q(p_i,q_i)\kappa_q(q_i) - p_i\kappa_q(p_i) + \xi(p_i,q_i)\Big)\,. \end{cases}$$

Here, ξ is the corrector introduced so that the Bregman- and interaction- quantities are synchronized. Indeed, then the individual quantities coincide, in particular,

$$\pi_q(p_i, q_i)\kappa_q(q_i) + \xi(p_i, q_i) = \sum_{i\in\mathbb{A}} \phi_q(p_i, q_i) \,.$$

Note that the corrector is independent of q. When seeking further physically founded explanations for the whole set-up it may well be important to take the corrector into account.

Quantities written out:

$$\Phi(P,Q) = \frac{1}{1-q}\Big(-1 + \sum_{i\in\mathbb{A}} \big(qp_i q_i^{q-1} + (1-q)q_i^q\big)\Big)\,,$$

$$\mathsf{S}(P) = \frac{1}{1-q}\Big(-1 + \sum_{i\in\mathbb{A}} p_i^q\Big)\,,$$

$$\mathsf{D}(P,Q) = \frac{1}{1-q}\sum_{i\in\mathbb{A}} \Big(qp_i q_i^{q-1} - p_i^q + (1-q)q_i^q\Big)\,.$$

Relaxing the condition of consistence: If we only assume that π is sound, i.e. that $\pi(x, x) = x$ for $0 \leq x \leq 1$, then other forms of interaction may leed to Tsallis-entropy as well. This happens with

$$\pi(x, y) = x^q y^{1-q} .$$

Thus, many quite different forms of interaction may give the same entropy function. But of course, the complexity- and divergence-functions will be different.

References in brief:

- Havrda and Charvát (1967): first appearence in the mathematical literature
- Lindhard and Nielsen (1971) and Lindhard (1974): first appearence in the physical literature
- Tsallis (1988): well known (:-)) take-off point which triggered much research and debate.

As recent contributions relevant for the present research, I mention Naudts (2008) and my own contribution from (2007).