CBPF-NF-033/94

# On Probabilities and Information – The Envelope Game

*Constantino Tsallis*[1], *Guy Deutscher*[2] *and Roger Maynard*[3]

[1]Centro Brasileiro de Pesquisas Físicas - CBPF
Rua Dr. Xavier Sigaud, 150
22.290-180 Rio de Janeiro/RJ, Brazil

[2]School of Physics and Astronomy,
University of Tel Aviv, Tel Aviv, Israel

[3]Laboratoire d'Expérimentation Numérique
Maison des Magistères - CNRS
B.P. 166, 38042 Grenoble Cedex 9, France

Abstract

By elaborating on the concept of information (herein represented by a specific generalization of the entropy) we propose an unified picture for statistical inference which treats on equivalent footing the "Bayesian" and "frequentist" views on the concept of probability. We illustrate on the Envelope Game.

**Key-words:** Generalized entropy; Statistical inference; Probabilities; Envelope game.

# 1 Introduction

The concepts of information, entropy, probabilities, statistical inference are deep and delicate, hence subject to multiple interpretations and controversies. Due to this fact, two main schools of thought have nowadays developed, namely the so called "Frequentist" and "Bayesian" ones [1]. The frequentists put the emphasis on the *limit of a frequency of (real or virtual) occurrence* as the correct definition of a probability; the Bayesians put the emphasis onto Bayes theorem. A very interesting discussion of relevant issues has been provided by Rodriguez [1] by using, as an illustration, the dilemme frequently referred to as the "Envelope Game". Let us recall this paradox in a quite essential version. Two similar envelopes are placed in front of you and you are informed that one of them contains the *double* of the money contained in the other one. You are authorized to definitively keep *only one* of the envelopes (and we assume you would like it to be the one which contains more money!). Once you choose one of them and before you open it, you are allowed to permutate it with the other envelope, and this as many times as you would like to. Is it your interest to do that?

Let us analyze the "naive" statistical inference reasoning: "This envelope contains a quantity of money $x$; consequently, the other one contains either $2x$ or $x/2$ with equal chances. Therefore, the expectation value is $(2x)(1/2) + (x/2)(1/2) = (5/4)x$. Since this value is *greater* than $x$ (which I have in hands), it is my interest to exchange it for the other". The argument still holds once you have exchanged, consequently your "rational" attitude should be an *endless sequence of permutations*, a quite absurd behavior! Clearly, the only way out of the paradox is to have an "expectation value" which *equals* $x$. Through a Bayesian analysis, Rodriguez exhibits the admissibility of the solution $(2x)(1/3) + (x/2)(2/3) = x$. Barron [2] provides basically the same solution along equivalent (Bayesian) lines by introducing the "sponsor" of the game (he simultaneously address an equivalent problem, namely the so called "Money Pump").

Furthermore, Rodriguez makes an analogy between Modern Physics and Statistical Inference. He states that, in the same manner we have in Physics,

$$\lim_{c\to\infty} \text{(Relativistic Mechanics)=(Classical Mechanics)} \tag{1}$$
$$(c \equiv \text{vacuum light velocity})$$

and

$$\lim_{h\to 0} \text{(Quantum Mechanics)=(Classical Mechanics)} \tag{2}$$
$$(h \equiv \text{Planck constant}),$$

we have, in Statistical Inference, that

$$\lim_{I\to I_0} \text{(Bayesian)=(Frequentist)} \tag{3}$$

where $I$ represents *prior information* and $I_0$ the state of knowledge of "Total Ignorance", i.e., absence of prior information.

It is the aim of the present note to suggest a more comprehensive frame in which *the concept of "probability" would be intimately related to the concept of "information"*. In other words, a given choice of "information" (herein represented by a specific generalization $S_q$ [3] of the Boltzmann-Gibbs-Shannon entropy) would be consistent with one and only one choice of "probability". If this idea was correct, Bayesian and Frequentist inferences would be treated on equal footing, and Eq. (3) would require further qualification (this point will become clear later on). The idea is illustrated with the Envelope Game.

# 2 Generalized Statistical Mechanics and Thermodynamics

It has been recently been proposed the generalization of Statistical Mechanics [3] and Thermodynamics [4] by using the following generalized entropy (in units of a conventional constant $k > 0$)

$$S_q = \frac{1 - \sum_i p_i^q}{q-1} \tag{4}$$

$$= -\sum_i p_i^q \frac{p_i^{1-q} - 1}{1-q} \qquad (q \in I\!R) \tag{4'}$$

(discrete version [3]; $\{p_i\}$ are the probabilities; $\sum_i p_i = 1$). This entropy yields the following related forms:

$$S_q = \frac{1 - C_q \int dx [p(x)]^q}{q-1} \tag{5}$$

$$= -\int dx [p(x)]^q \frac{[p(x)]^{1-q} - C_q}{1-q} \tag{5'}$$

(continuous version [5]; $C_q$ is a constant whose dimension is that of $[x]^{q-1}$, and which is irrelevant for the entropy optimization problem; if $x$ is a pure number, we can take $C_q = 1$)

$$S_q = \frac{1 - Tr\rho^q}{q-1} \tag{6}$$

$$= -Tr\rho^q \frac{\rho^{1-q} - 1}{1-q} \tag{6'}$$

(off-diagonal version [6], $\rho$ being the density operator)

$$\tilde{S}_q = -\sum_i p_i \frac{\left(p_i/p_i^{(0)}\right)^{q-1} - 1}{q-1} \tag{7}$$

(discrete version of the cross or relative entropy [7]; $\{p_i^{(0)}\}$ are reference probabilities, frequently taken to be those corresponding to equilibrium; $\sum_i p_i^{(0)} = 1$) and

$$\tilde{S}_q = -\int dx p(x) \frac{[p(x)/p^{(0)}(x)]^{q-1} - 1}{q-1} \tag{8}$$

(continuous version of the cross or relative entropy [8]; if $p^{(0)}(x)$ independs from $x$, we can identify the constant $C_q$ appearing in Eq. (5) with $\left(p^{(0)}\right)^{1-q}$).

In the $q \to 1$ limit (and using $z^{1-q} \sim 1 + (1-q)\ln z$ for arbitrary positive $z$) we recover the standard expressions

$$S_1 = -\sum_i p_i \ln p_i \tag{9}$$

(Shanon entropy [9])

$$S_1 = -\int dx p(x) \ln p(x) + C \tag{10}$$

(Boltzmann-Gibbs entropy [10]; $C$ is a constant which depends on the dimension of $x$ and which is irrelevant for the entropy optimization problem; if $x$ is a pure number we can take $C = 0$)

$$S_1 = -Tr\rho \ln \rho \tag{11}$$

(von Neumann entropy [11])

$$\tilde{S}_1 = -\sum_i p_i \ln(p_i/p_i^{(0)}) \tag{12}$$

(discrete version of Kullback-Leibler [12] or Shannon-Jaynes [13] cross entropy)

and

$$\tilde{S}_1 = -\int dx p(x) \ln[p(x)/p^{(0)}(x)] \tag{13}$$

(continuous version of Kullback-Leibler or Shannon-Jaynes cross entropy; if $p^{(0)}$ independs from $x$, we can identify the constant $C$ appearing in Eq. (10) with $\ln(p^{(0)})$.

$S_q$ as given in Eq. (4) belongs to the large class of entropies considered in Information Theory [14], and exhibits the same functional form (but a different $q$-dependent prefactor) as the entropy introduced by Havrda and Charvat [15] and by Daroczy [16]. Futhermore, it is related the so called Renyi entropy [3, 16] $S_q^R \equiv (\ln \sum_i p_i^q)/(1-q)$ through

$$S_q^R = \frac{\ln[1 + (1-q)S_q]}{1-q} \tag{14}$$

$S_q$ is positive ($\forall q$), expansible (for $q > 0$), concave (for $q > 0$; it is convex for $q < 0$), invariant under permutations within the $\{p_i\}$ ($\forall q$), extremal for equiprobability (maximal for $q > 0$, and minimal for $q < 0$), a monotonically increasing function of the total number $W$ of possible configurations ($\forall q; \sum_{i=1}^{W} p_i = 1$), a decreasing function of $q$ for fixed $\{p_i\}$, continuous in $\{p_i\}$ (in the interval [0,1] for $q > 0$, and in the interval (0,1) for $q < 0$). If $A$ and $B$ are *independent systems* (in the sense that $p_{ij}^{A \cup B} = p_i^A p_j^B, \forall(i,j)$), $S_q$

is *pseudo-additive* (or *pseudo-extensive*) [17]:

$$S_q^{A \cup B} = S_q^A + S_q^B + (1-q) S_q^A S_q^B \tag{15}$$

Let us remark at this point that $S_q$ is *substantially* different from $S_q^R$. Indeed, it is generically *concave* and *nonextensive*, whereas $S_q^R$ is generically *nonconcave* and *extensive* ($S_q^{R(A \cup B)} = S_q^{R(A)} + S_q^{R(B)}$). This makes an enormous difference if *physical* applications are focused. Indeed, concavity guarantees the *thermodynamic stability* of the system (positive specific heat [17, 18], generalized fluctuation-dissipation theorem [19], etc.), whereas extensivity can be seen as a mathematically simplifying *linearity* property which might be present or not. Before going on to further physical considerations, let us mention a remarkable property [4] (which generalizes the celebrated Shannon additive property). If we define $p_L \equiv \sum_{i=1}^{W_L}$ and $p_M \equiv \sum_{i=W_L+1}^{W} p_i \quad (p_L + p_M = 1)$, we easily prove that

$$S_q(\{p_i\}) = S_q(p_L, p_M) + p_L^q S_q(\{p_i/p_L\}) + p_M^q S_q(\{p_i/p_M\}) \tag{16}$$

where the sets $\{p_i\}, \{p_i/p_L\}$ and $\{p_i/p_M\}$ contain respectively $W$, $W_L$ and $(W - W_L)$ probabilities.

The likelihood function associated with $\{p_i\}$ is given by [19, 20]

$$P_q(\{p_i\}) \propto [1 + (1-q) S_q(\{p_i\})]^{\frac{1}{1-q}} \tag{17}$$

In the $q \to 1$ limit we recover the Einstein expression [21] $P_1(\{p_i\}) \propto e^{S_1(\{p_i\})}$. To obtain the equilibrium distribution we must extremize $P_q(\{p_i\})$, hence $S_q(\{p_i\})$, under the given constraints. For understanding how the generalization of Statistical Mechanics proceeds, it suffices to discuss the *grand-canonical ensemble*. The equilibrium distribution $\{p_i^{eq}\}$ optimizes $S_q$ under the constraints

$$\sum_i p_i = 1 \tag{18}$$

$$< \mathcal{H} >_q \equiv \sum_i p_i^q \varepsilon_i = U_q \tag{19}$$

($\{\varepsilon_i\} \equiv$ set of eigenvalues of the Hamiltonian $\mathcal{H}$; $U_q$ is assumed finite and known)

$$< O^{(m)} >_q \equiv \sum_i p_i^q O_i^{(m)} = O_q^{(m)} \quad (m = 1, 2, ...) \tag{20}$$

($O_i^{(m)}$ are the eigenvalues of the observable $O^{(m)}$, herein assumed, for simplicity, to commutate with $\mathcal{H}$; $\{O_q^{(m)}\}$ are assumed finite and known). It is easily proved that [3, 6]

$$p_i^{eq} = \frac{\left\{1 - \beta(1-q)[\varepsilon_i + \sum_m \mu_m O_i^{(m)}]\right\}^{\frac{1}{1-q}}}{Z_q} \tag{21}$$

with the generalized *grand-partition function* given by

$$Z_q \equiv \sum_i \{1 - \beta(1-q)[\varepsilon_i + \sum_m \mu_m O_i^{(m)}]\}^{\frac{1}{1-q}} \tag{22}$$

$\beta$ and $\{\beta\mu_m\}$ being the Lagrange parameters. Eqs. (21) and (22) recover, in the $q \to 1$ limit, the standard (Boltzmann-Gibbs) exponential law. The *canonical ensemble* is obtained when the constraints (20) are not present, hence $\mu_m = 0$ ($\forall m$), hence

$$p_i^{eq} = [1 - \beta(1-q)\varepsilon_i]^{\frac{1}{1-q}} / Z_q \tag{23}$$

with

$$Z_q \equiv \sum_i [1 - \beta(1-q)\varepsilon_i]^{\frac{1}{1-q}} \tag{24}$$

The *microcanonical ensemble* is obtained when even the constraint (19) is absent, hence $\beta = 0$, hence

$$p_i^{eq} = 1/W \quad (\forall i) \tag{25}$$

It can be also proved in general that (using $\beta \equiv 1/T$) [4, 6]

$$\frac{1}{T} = \frac{\partial S_q}{\partial U_q} \tag{26}$$

$$\frac{\mu_m}{T} = \frac{\partial S_q}{\partial O_q^{(m)}} \quad (\forall m) \tag{27}$$

$$U_q = -\frac{\partial}{\partial \beta} \frac{Z_q^{1-q} - 1}{1-q} \tag{28}$$

$$O_q^{(m)} = -\frac{\partial}{\partial(\beta\mu_m)} \frac{Z_q^{1-q} - 1}{1-q} \tag{29}$$

and

$$F_q \equiv U_q - TS_q - \sum_m \mu_m O_q^{(m)} \tag{30}$$
$$= -\frac{1}{\beta}\frac{Z_q^{1-q}-1}{1-q}$$

The presence of $p_i^q$ (instead of $p_i$) in the constraints (19) and (20) demands some comments. Surprising at first sight, it is in fact very natural since: (i) it appears in Eq. (16); (ii) it enables a connection with appropriately generalized (nonextensive) Thermodynamics which preserves the standard Legendre transform framework [4] (or equivalently, it automatically satisfies Jaynes Information Theory duality relations [6]); (iii) it produces, at long distances, *power*-law distributions (instead of *exponential* ones) within an entropic variational procedure which uses acceptable a priori constraints [22]; indeed, if we optimize $S_q$ with the constraints $\int dx p(x) = 1$ and $< x^2 >_q \equiv \int dx [p(x)]^q x^2 = \sigma_q^2$ ($\sigma_q$ is *finite* and known) we obtain $p(x) \propto 1/[1+\beta(q-1)x^2]^{\frac{1}{q-1}}$, which recovers the Gaussian law for $q \to 1$, the Lorentzian law for $q = 2$, and generic *power*-laws $p(x) \propto 1/x^{\frac{2}{q-1}}$ if $x \to \infty$ and $q \neq 1$ (the benefic effect of using $p^q$ can be seen, for instance, for $q = 2$: for the Lorentzian distribution $p(x) \propto 1/(1+\beta x^2)$, $< x^2 >_2$ is *finite* whereas $< x^2 >_1 \equiv \int dx x^2 p(x)$ *diverges*; (iv) it produces $q$-invariant forms for the Ehrenfest theorem [6], the von Neumann equation [23], one of the Einstein relations for stochastic movement [24], among other remarkable properties. The *q-expectation value* $< O >_q \equiv \sum_i p_i^q O_i$ (associated with the observable $O$) can be interpreted as a standard mean value of $(p_i^{q-1} O_i)$, but *should by no means be interpreted, for $q \neq 1$, as a standard mean value of $O$*. Indeed, $< \lambda >_q \neq \lambda$ for any nonvanishing constant $\lambda$.

# 3 Information, Probabilities and the Envelope Game

Let us now address a central point: what is the meaning of $q$ in terms of inference? Jumarie [25] argued, for the Renyi entropy, that $q < 1, q = 1$ and $q > 1$ correspond respectively to *prior knowledge, no prior knowledge* and *prior misknowledge.* Since $S_q$ is, for fixed $q$, a monotonically increasing function of $S_q^R$ (see Eq. (14)), the same interpretation should apply to $S_q$. Let us illustrate this fact through a simple example. At $t = 0$, *four* boxes are presented to you, and you are informed that one (and only one) of them contains a chocolate. Since for equiprobability ($p_i = 1/W, \forall i$) we have $S_q = (W^{1-q} - 1)/(1-q)$ (generalization of $S_1 = \ln W$), then $S_0(t=0) = 3, S_1(t=0) = \ln 4$ and $S_2(t=0) = 3/4$. At $t = 1$, you are informed that the chocolate is in one of *two* specific boxes (among the four ones). Then $S_0(t=1) = 1, S_1(t=1) = \ln 2$ and $S_2(t=1) = 1/2$. Finally, at $t = 2$, the good box is indicated. Consequently, $S_0(t=2) = S_1(t=2) = S_2(t=2) = 0$. See Fig. 1, where the relevant quantity $S_q(t)/S_q(0)$ is represented. We verify that the time-evolution towards full knowledge (i.e., $S_q = 0$) is *slower* for increasingly large $q$. Since $q = 1$ is known to correspond to *no prior knowledge* (see, for instance, the "kangaroo argument" [26]), the association of *prior knowledge (prior misknowledge)* to $q < 1 (q > 1)$ is intuitively consistent.

Let us finally address the Envelope Game. *We suggest here that the "expectation value" which drives the decision is the q-expectation value* $<gain>_q$. For the envelope in hands it is $(x)1^q = x$. For the other envelope it is $(2x)p^q + (x/2)(1-p)^q$. As already argued, it must be $< gain >_q^{1^{st} envelope} = < gain >_q^{2^{nd} envelope}$, hence

$$2p^q + \frac{1}{2}(1-p)^q = 1 \tag{31}$$

which is represented in Fig. 2. The Bayesian answer $p = 1/3$ (see [1, 2]) corresponds to $q = 1$ information (no prior information), whereas the frequentist answer $p = 1/2$ corresponds to $q = \ln(5/2)/\ln 2$ ($\simeq 1.322$) information (prior misknowledge). *They are equivalent among them and equivalent to any other solution of Eq.* (31). If we go back now to physical analogies, the situation is fully analogous to Quantum Mechanics. Indeed, within this formalism, the time-evolution of the physical quantities can be *exclusively associated with the wave functions* (Schroedinger representation), or *exclusively associated with the operators* (Dirac representation) or *partially with both wave functions and operators* (Heisenberg or mixed representation). In terms of schematic formulas (like Eqs.

(1-3)), we would have

$$\text{(Bayesian, } q = 1 \text{ information)=(Frequentist, } q\text{-information)} \qquad (32)$$

thus recovering, in particular, the well known result that Bayesian and Frequentist analysis coincide *whenever no prior knowledge is available* (i.e., $q \to 1$). Also, Rodriguez "formula" (present Eq. (3)) would be interpreted as follows:

$$\begin{aligned} &\lim_{I \to I_0} \text{(Bayesian, q=1 information)} \\ &= \lim_{q \to 1} \text{(Frequentist, q-information)} \\ &= \text{(Frequentist, q=1 information)} \end{aligned} \qquad (33)$$

where we have used Eq. (32).

Before ending, it is instructive to consider the generalization of the Envelope Game in the sense that one of the envelopes contains (instead of the *double)* $\rho$ times the money contained in the other one ($\rho \geq 0$). Eq. (31) is generalized into

$$\rho p^q + \frac{1}{\rho}(1-p)^q = 1, \qquad (34)$$

hence $p(q,\rho) + p(q, 1/\rho) = 1$.

The Bayesian viewpoint corresponds to $q = 1$; consequently Eq. (34) implies

$$p_B = \frac{1}{\rho + 1}, \qquad (35)$$

hence $p_B(\rho) + p_B(1/\rho) = 1$ (see Fig. 3).

The Frequentist viewpoint corresponds to $p = 1/2$; consequently Eq. (34) implies

$$q_F = \frac{\ln(\rho + \frac{1}{\rho})}{\ln 2} \qquad (36)$$

hence $q_F(\rho) = q_F(1/\rho)$ (see Fig. 4).

The $\rho \to 1$ limit is an interesting situation. Eq. (34) implies

$$p^q + (1-p)^q = 1 \qquad (37)$$

whose unique solutions are either $(q = 1, \forall p)$ or $(q \neq 1, p = 0$ or $1)$; furthermore, the Bayesian point $(p, q) = (p_B, 1)$ collapses with the Frequentist one $(1/2, q_F)$. The global situation is depicted in Fig. 5.

Further statistical inference examples in order to test the present proposal would be very welcome.

A satisfactory outcome of the present discussion would certainly enlighten Equilibrium and Nonequilibrium Statistical Mechanics, where confrontation between the two schools does exist. Dougherty [27] has carefully compared both philosophies. He refers to the frequentists as the *Brussels school* ("objective probability"; relevant names: L. Boltzmann, P. Ehrenfest, R.C. Tolman, L. Landau, S. Chapman, D. Enskog, N.N. Bogoliubov, D.N. Zubarev, I. Prigogine, R. Balescu, J.L. Lebowitz), which uses, as starting point, the *microscopic dynamics*. He refers to the Bayesians as the *Maxent school* ("subjective probability", "degree of belief", "information theory", and other anthropomorfic concepts; relevant names: J.W. Gibbs, C.E. Shannon, E.P. Wigner, E.T. Jaynes, W.T. Grandy, A.J.M. Garrett, D.N. Zubarev, R. Balian), which uses, as starting point, the *optimization of an appropriate entropy* with conveniently chosen constraints. To reinforce the proposal for *unification* that we have developed in the present paper, let us quote some of Dougherty's statements: "*···thus revealing the equivalence of the methods of the Brussels school and the Maxent school···*", "*As we believe that the work of the two schools is in the end equivalent we can adopt a unified attitude to the question*", "*··· my conclusion is that, not withstanding the tacit antagonism between the two schools, the different-looking formalisms and the differing fields of application, they are very likely to be equivalent*", "*Eventually it should be possible to offer a broader formalism that makes the best use of the mathematical techniques of both.*", "*If the two formalisms are indeed equivalent, it may be possible to weld them together, or to present them as special cases or aspects of something more general.*", "*··· they can be seen as part of a common underlying structure.*" Also, Behara ([14], page 19) writes (unfortunately with no further details): "*During the past few years, researchers in information theory have been busy in investigating the relationships* *between the notions of* **probability** *and* **information**. *We do not go into these details here. However, it would be appropriate to mention that* **probability** *of occurrence of an event can also be regarded as a function of the amount of* **information** *yielded by that event."* By introducing a generalized entropy $S_q$ within the methodology of *Information Theory* (Bayesian philosophy), and in principle determining (as suggested by the results presented in [5] and [22]), through the particular *microscopic dynamics* (frequentist philosophy), the correct value of $q$ to be used, we believe that we are exhibiting a realization of Dougherty's philosophical program and of Behara's comment. To close these general remarks, let us add an analogy with the historical evolution of the concept of "motion". For most pre-socratic philosophers, motion was an *intrinsic* property of the moving object; Galileo showed that this is essentially wrong since motion *must* be thought with reference to an external frame. Analogously, we are herein suggesting that probability *must* be thought with reference to an information frame (implicitely or explicitely taken into account in the experimental protocol of the phenomenon we are focusing).

It is our pleasure to acknowledge J. Schweitzer, who draw our attention onto the Envelope Game and related references. Two of us (C.T. and G.D.) are grateful for warm hospitality received at the Laboratoire d'Expérimentation Numérique (Université Joseph Fourier de Grenoble). Also, computational assistance by J.A. Redinz is acknowledged.

# Caption for Figures

**Fig. 1** - Time evolution of the $q = 0, 1, 2, \infty$ entropies for the "chocolate in the boxes" problem.

**Fig. 2** - Locus in the $(q, p)$ plane where the paradox does not exist ($q \equiv$entropy index; $p \equiv$probability of having the envelope which contains more money). The point corresponding to the "naive" reasoning which leads to the paradox, as well as those corresponding to the Bayesian and Frequentist solutions, are indicated.

**Fig. 3** - The Bayesian solution when one envelope contains $\rho$ times the money contained in the other one.

**Fig. 4** - The Frequentist solution when one envelope contains $\rho$ times the money contained in the other one.

**Fig. 5** - The general solution (in the $(q, p)$ plane, for typical values of $\rho$) when one envelope contains $\rho$ times the money contained in the other one. $\rho = 2$ reproduces Fig. 2; $\rho = 0$ yields $p = 1, \forall q < \infty$; $\rho \to \infty$ yields $p = 0, \forall q < \infty$.

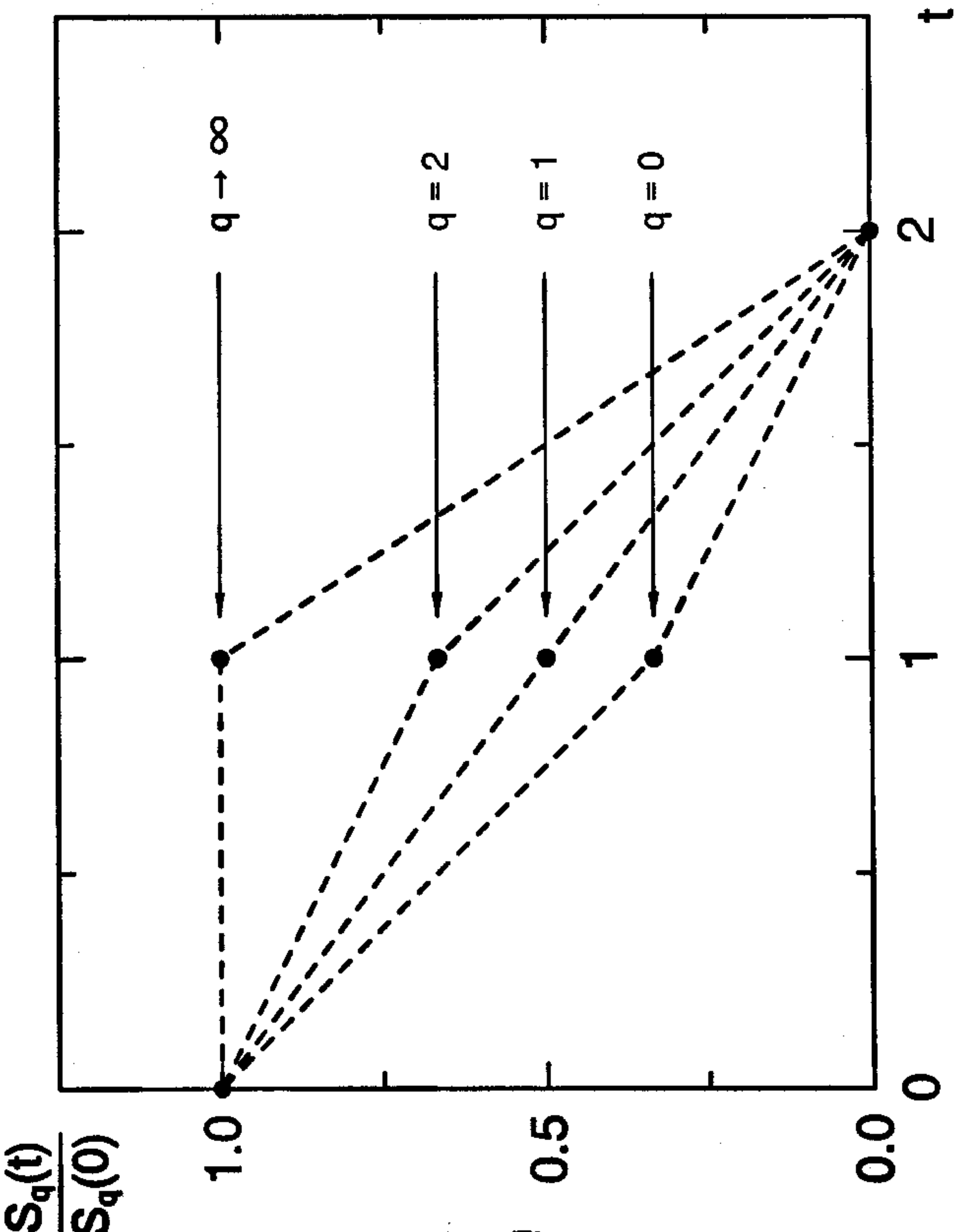


Fig. 1

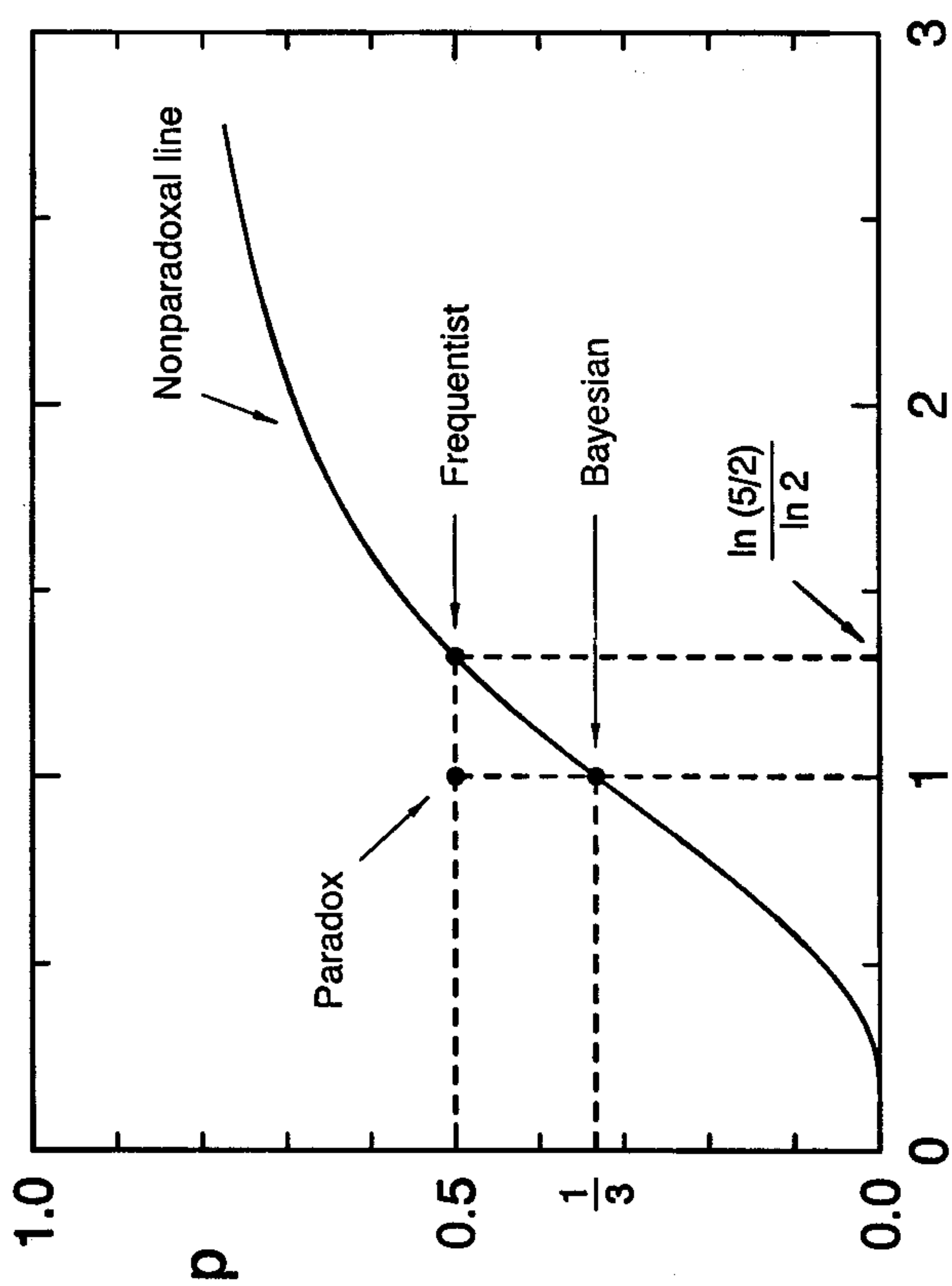
Nonparadoxal line
Frequentist
Bayesian
Paradox
$\frac{\ln(5/2)}{\ln 2}$
q
0
1
2
3
p
1.0
0.5
$\frac{1}{3}$
0.0

Fig. 2

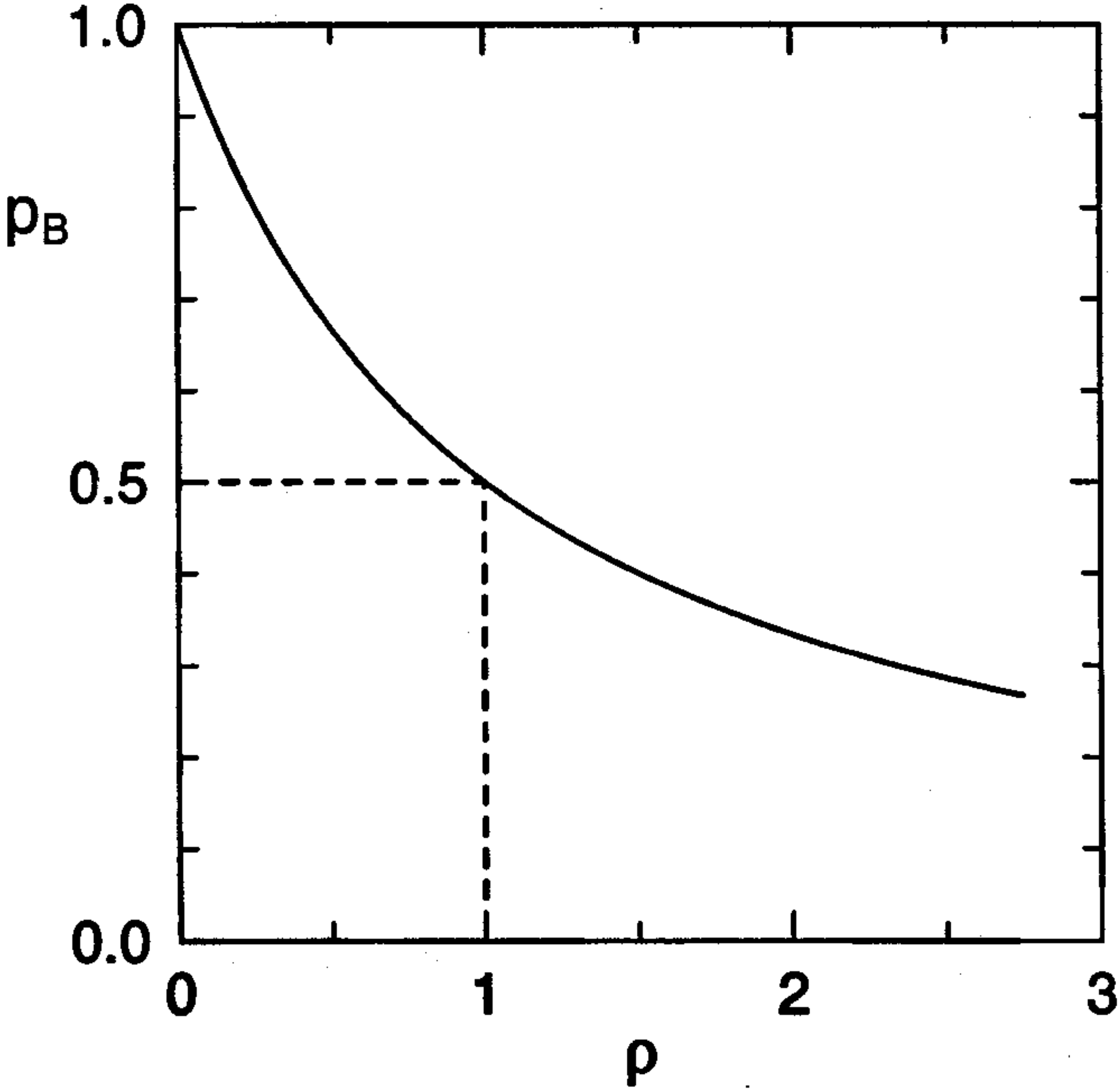


Fig. 3

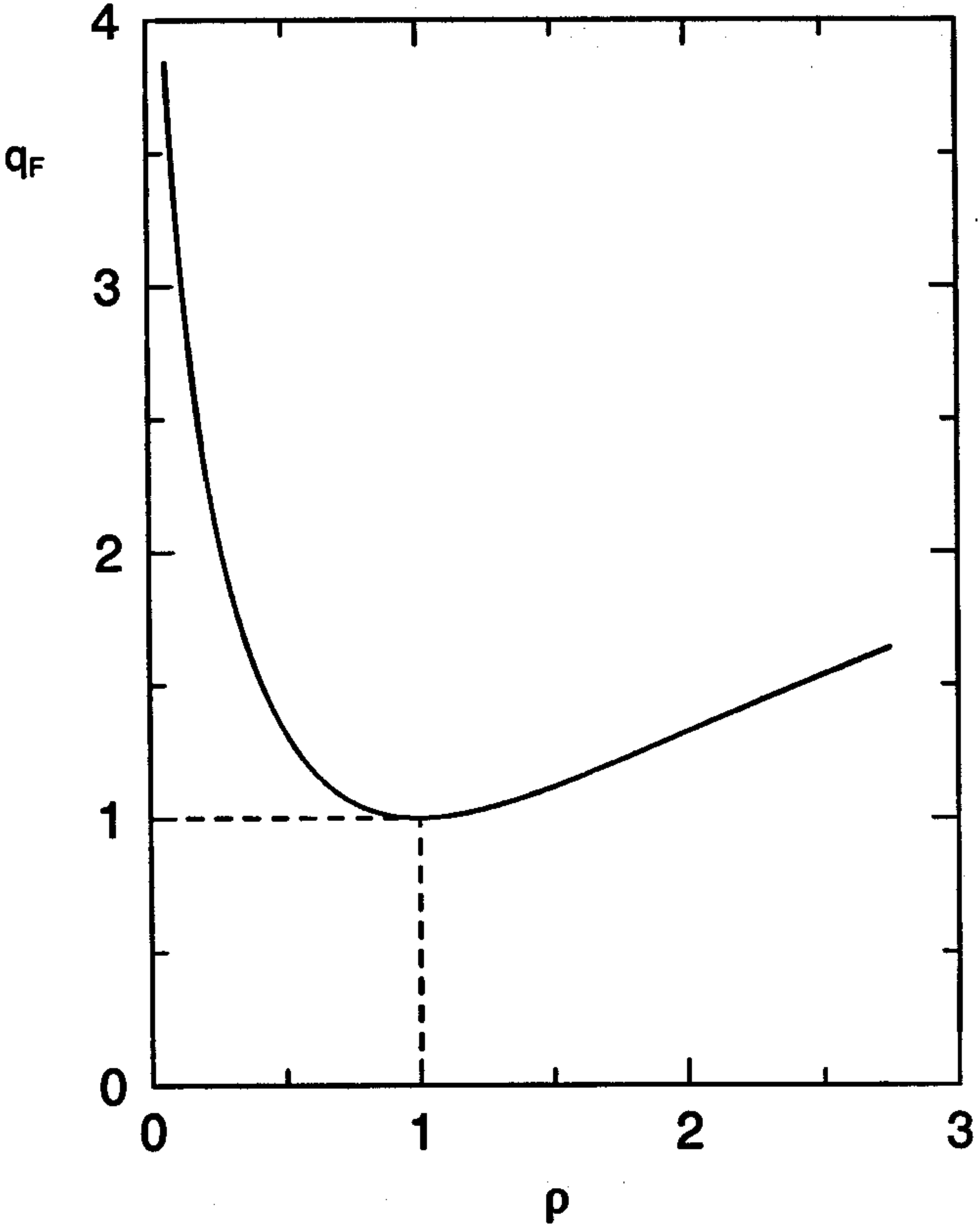


Fig. 4

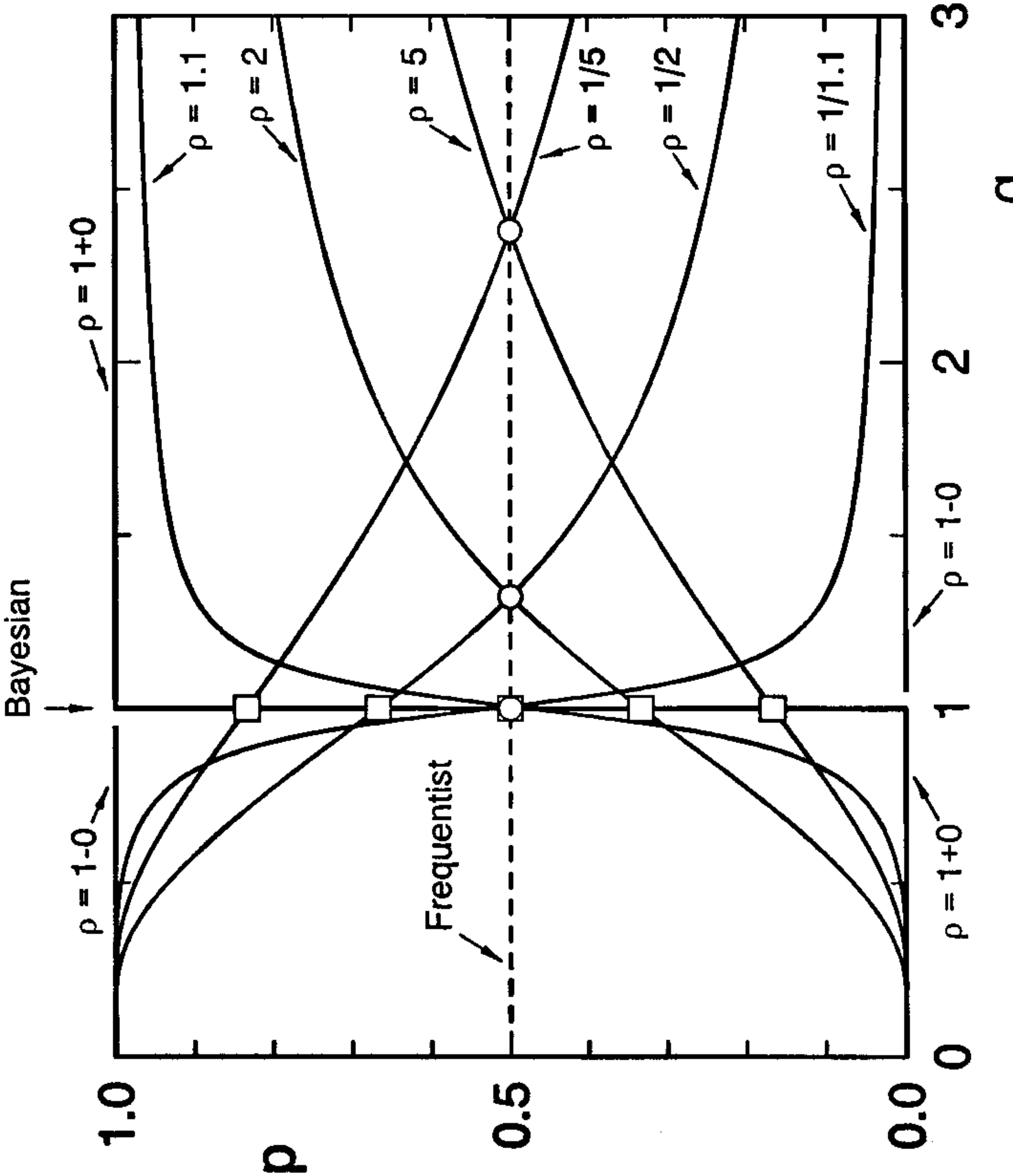
Bayesian
Frequentist
ρ = 1+0
ρ = 1-0
ρ = 1.1
ρ = 2
ρ = 5
ρ = 1/5
ρ = 1/2
ρ = 1/1.1
ρ = 1-0
ρ = 1+0
p
1.0
0.5
0.0
q
0
1
2
3

Fig. 5

# References

[1] C.C. Rodriguez, in "Maximum Entropy and Bayesian Methods in Science and Engineering", ed. G.J. Erickson and C.R. Smith (Kluwer Academic, Dordrecht, 1988), vol. I, p. 189.

[2] R. Barron, in "Maximum Entropy and Bayesian Methods", ed. J. Skilling (Kluwer Academic, Dordrecht, 1989), p. 423.

[3] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).

[4] E.M.F. Curado and C. Tsallis, J. Phys. A **24**, L69 (1991); Corrigenda: J. Phys. A **24**, 3187 (1991) and **25**, 1019 (1992).

[5] A.R. Plastino and A. Plastino, Phys. Lett. A **174**, 384 (1993); J.J. Aly, in "N-Body Problems and Gravitational Dynamics", Proceedings of the Meeting held at Aussois-France (21-25 March 1993), eds. F. Combes and E. Athanassoula (Publications de l'Observatoire de Paris) p.19.

[6] A.R. Plastino and A. Plastino, Phys. Lett. A **177**, 177 (1993).

[7] J.D. Ramshaw, Phys. Lett. A **175**, 169 (1993).

[8] C. Tsallis, "Generalized entropy-based criterion for consistent nonparametric testing", preprint (1994).

[9] C.E. Shannon, "A mathematical theory of communication", Bell System Technical Journal **27**, 379 and 623 (1948).

[10] L. Boltzmann, Wien, Ber. **66**, 275 (1872) and Ann. Phys. (Leipzig) **57**, 773 (1896) [translated and reprinted in S.G. Brush, "Kinetic Theory 2" (Pergamon, New York, 1966)]; J.W. Gibbs, "Elementary Principles in Statistical Mechanics" (Yale University Press, New Haven, 1902) [reprinted by Dover, New York, 1960].

[11] J. von Neumann, "Mathematical Foundations of Quantum Mechanics" (Princeton University Press, Princeton, 1955).

[12] S.L. Kullback and R.A. Leibler, Ann. Stat. **22**, 79 (1951); S.L. Kullback, in "Information Theory and Statistics" (J. Wiley, New York, 1959).

[13] E.T. Jaynes, in R.D. Rosenkrantz, ed. "E.T. Jaynes: Papers on Probability, Statistics and Statistical Physics" (D. Reidel, Dordrecht, 1983).

[14] J.C.A. van der Lubbe, Y. Boxma and D.E. Boekee, Inform. Sci. **32**, 187 (1984); D.E. Boekee and J.C.A. van der Lubbe, Information and Control **45**, 136 (1980); M. Behara, "Additive and Nonadditive Measures of Entropy" (Wiley Eastern, New Delhi, 1990).

[15] J. Havrda and F. Charvat, Kybernetica **3**, 30 (1967); Z. Daroczy, Information and Control **16**, 36 (1970).

[16] A. Renyi, "On measures of entropy and information", in Proc. Fourth Berkeley Symposium on Mathematical Statistics and Probability **1**, 547 (1960), and "Probability Theory" (North-Holland, Amsterdam, 1970); according to I. Csiszar ["Information measures: A critical survey", in "Transactions of the Seventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes" and the European Meeting of Statisticians, 1974 (Reidel, Dordrecht, 1978), p. 73], this entropy was introduced by Schützenberger.

[17] C. Tsallis, "Extensive versus Nonextensive Physics" in "New Trends in Magnetic Materials and their Applications", ed. J.L. Morán-Lopez and J.M. Sanchez (Plenum Press, New York, 1994), in press.

[18] E.P. da Silva, C. Tsallis and E.M.F. Curado, Physica A **199**, 137 (1993); errata: Physica A **203**, 160 (1994).

[19] A. Chame and E.V.L. de Mello, J. Phys. A **27**, 3663 (1994).

[20] C. Tsallis, "Some Comments on Boltzmann-Gibbs Statistical Mechanics", in "Chaos, Solitons and Fractals", ed. G. Marshall (Pergamon Press, Oxford, 1994), in press.

[21] A. Einstein, Ann. Phys. **33**, 1275 (1910) and "Investigations on the Theory of the Brownian Movement" (Methuen and Co, London, 1926); H.B. Callen, "Thermodynamics" (J. Wiley, New York, 1960) chapter 15, p. 269.

[22] P.A. Alemany and D.H. Zanette, Phys. Rev. E **49**, 956 (1994).

[23] A.R. Plastino and A. Plastino, Physica A **202**, 438 (1994).

[24] D.A. Stariolo, Phys. Lett. A **185**, 262 (1994).

[25] G. Jumarie, "Relative Information. Theories and Applications" (Springer-Verlag, Berlin, 1988); Cybernetics and Systems **19**, 169 and 311 (1988).

[26] S.F. Gull and J. Skilling, IEE Proceedings **131**, part F (1984) p. 646; D.S. Sivia, Los Alamos Science **19**, 180 (1990).

[27] J.P. Dougherty, Stud. Hist. Phil. Sci. **24**, 843 (1993) and Phil. Trans. R. Soc. Lond. A **346**, 259 (1994).