STATISTICAL INDIAN INSTITUTE

UNITY IN DIVERSITY

Platinum Jubilee Series

Statistical Science and Interdisciplinary Research — Vol. 7

Perspectives in Mathematical Sciences I

Probability and Statistics

Editors

N. S. Narasimha Sastry

T. S. S. R. K. Rao

Mohan Delampady

B. Rajeev

Series Editor: ***Sankar K. Pal***

World Scientific

Platinum Jubilee Series

Statistical Science and
Interdisciplinary Research — Vol. 7

Perspectives in Mathematical Sciences I

Probability and Statistics

Editors

N. S. Narasimha Sastry

T. S. S. R. K. Rao

Mohan Delampady

B. Rajeev

Indian Statistical Institute, India

Series Editor: ***Sankar K. Pal***

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Statistical Science and Interdisciplinary Research — Vol. 7
PERSPECTIVES IN MATHEMATICAL SCIENCES I
Probability and Statistics

ISBN-13 978-981-4273-62-6
ISBN-10 981-4273-62-7

Printed in Singapore by World Scientific Printers

Foreword

The Indian Statistical Institute (ISI) was established on 17th December, 1931 by a great visionary Professor Prasanta Chandra Mahalanobis to promote research in the theory and applications of statistics as a new scientific discipline in India. In 1959, Pandit Jawaharlal Nehru, the then Prime Minister of India introduced the ISI Act in the parliament and designated it as an *Institution of National Importance* because of its remarkable achievements in statistical work as well as its contribution to economic planning.

Today, the Indian Statistical Institute occupies a prestigious position in the academic firmament. It has been a haven for bright and talented academics working in a number of disciplines. Its research faculty has done India proud in the arenas of Statistics, Mathematics, Economics, Computer Science, among others. Over seventy five years, it has grown into a massive banyan tree, like the institute emblem. The Institute now serves the nation as a unified and monolithic organization from different places, namely Kolkata, the Headquarters, Delhi, Bangalore, and Chennai, three centers, a network of five SQC-OR Units located at Mumbai, Pune, Baroda, Hyderabad and Coimbatore, and a branch (field station) at Giridih.

The platinum jubilee celebrations of ISI have been launched by Honorable Prime Minister Prof. Manmohan Singh on December 24, 2006, and the Government of India has declared 29th June as the "Statistics Day" to commemorate the birthday of Professor Mahalanobis nationally.

Professor Mahalanobis, was a great believer in interdisciplinary research, because he thought that this will promote the development of not only Statistics, but also the other natural and social sciences. To promote interdisciplinary research, major strides were made in the areas of computer science, statistical quality control, economics, biological and social sciences, physical and earth sciences.

The Institute's motto of "unity in diversity" has been the guiding principle of all its activities since its inception. It highlights the unifying role of statistics in relation to various scientific activities.

In tune with this hallowed tradition, a comprehensive academic programme, involving Nobel Laureates, Fellows of the Royal Society, Abel prize winner and other dignitaries, has been implemented throughout the Platinum Jubilee year, highlighting the emerging areas of ongoing frontline research in its various scientific divisions, centers, and outlying units. It includes international and national-level seminars, symposia, conferences and workshops, as well as series of special lectures. As an outcome of these events, the Institute is bringing out a series of comprehensive volumes in different subjects under the title *Statistical Science and Interdisciplinary Research*, published by the World Scientific Press, Singapore.

The present volume titled *Perspectives in Mathematical Sciences I: Probability and Statistics* is the seventh one in the series. The volume consists of eleven chapters, written by eminent probabilists and statisticians from different parts of the world. These chapters provide a current perspective of different areas of research, emphasizing the major challenging issues. They deal mainly with statistical inference, both frequentist and Bayesian, with applications of the methodology that will be of use to practitioners. I believe the state-of-the art studies presented in this book will be very useful to both researchers as well as practitioners.

Thanks to the contributors for their excellent research contributions, and to the volume editors Profs. N. S. Narasimha Sastry, T. S. S. R. K. Rao, M. Delampady and B. Rajeev for their sincere effort in bringing out the volume nicely in time. Initial design of the cover by Mr. Indranil Dutta is acknowledged. Sincere efforts by Prof. Dilip Saha and Dr. Barun Mukhopadhyay for editorial assistance are appreciated. Thanks are also due to World Scientific for their initiative in publishing the series and being a part of the Platinum Jubilee endeavor of the Institute.

December 2008
Kolkata

Sankar K. Pal
Series Editor and
Director

Preface

Indian Statistical Institute, a premier research institute founded by Professor Prasanta Chandra Mahalanobis in Calcutta in 1931, celebrated its platinum jubilee during the year 2006-07. On this occasion, the institute organized several conferences and symposia in various scientific disciplines in which the institute has been active.

From the beginning, research and training in probability, statistics and related mathematical areas including mathematical computing have been some of the main activities of the institute. Over the years, the contributions from the scientists of the institute have had a major impact on these areas.

As a part of these celebrations, the Division of Theoretical Statistics and Mathematics of the institute decided to invite distinguished mathematical scientists to contribute articles, giving "a perspective of their discipline, emphasizing the current major issues". A conference entitled "Perspectives in Mathematical Sciences" was also organized at the Bangalore Centre of the institute during February 4-8, 2008.

The articles submitted by the speakers at the conference, along with the invited articles, are brought together here in two volumes (Part I and Part II).

Part I consists of articles in Probability and Statistics. Articles in Statistics are mainly on statistical inference, both frequentist and Bayesian, for problems of current interest. These articles also contain applications illustrating the methodologies discussed. The articles on probability are based on different "probability models" arising in various contexts (machine learning, quantum probability, probability measures on Lie groups, economic phenomena modelled on iterated random systems, "measure free martingales", and interacting particle systems) and represent active areas of research in probability and related fields.

Part II consists of articles in Algebraic Geometry, Algebraic Number Theory, Functional Analysis and Operator Theory, Scattering Theory, von Neumann Algebras, Discrete Mathematics, Permutation Groups, Lie Theory and Super Symmetry.

All the authors have taken care to make their exposition fairly self-contained. It is our hope that these articles will be valuable to researchers at various levels.

The editorial committee thanks all the authors for writing the articles and sending them in time, the speakers at the conference for their talks and various scientists who have kindly refereed these articles. Thanks are also due to the National Board for Higher Mathematics, India, for providing partial support to the conference. Finally, we thank Ms. Asha Lata for her help in compiling these volumes.

October 16, 2008

N. S. Narasimha Sastry
T. S. S. R. K. Rao
Mohan Delampady
B. Rajeev

Contents

Chapter 1

Entropy and Martingale

K. B. Athreya[1] and M. G. Nadkarni[2]

[1] *Department of Mathematics,*
Iowa State University, Ames, Iowa, 50011, USA
and
I. M. I, Department of Mathematics,
Indian Institute of Science, Bangalore, 560012, India
kba@iastate.edu

[2] *Department of Mathematics*
University of Mumbai, Kalina, Mumbai, 400098, India
mgnadkarni@gmail.com

1.1. Introduction

This article discusses the concepts of relative entropy of a probability measure with respect to a dominating measure and that of measure free martingales. There is considerable literature on the concepts of relative entropy and standard martingales, both separately and on connection between the two. This paper draws from results established in [1] (unpublished notes) and [6]. In [1] the concept of relative entropy and its maximization subject to a finite as well as infinite number of linear constraints is discussed. In [6] the notion of measure free martingale of a sequence $\{f_n\}_{n=1}^{\infty}$ of real valued functions with the restriction that each f_n takes only finitely many distinct values is introduced. Here is an outline of the paper.

In section 1.2 the concepts of relative entropy and Gibbs-Boltzmann measures, and a few results on the maximization of relative entropy and the weak convergence of of the Gibbs-Boltzmann measures are presented. We also settle in the negative a problem posed in [6]. In section 1.3 the notion of measure free martingale is generalized from the case of finitely many valued sequence $\{f_n\}_{n=1}^{\infty}$ to the general case where each f_n is allowed to be a Borel function taking possibly uncountably many values. It is shown that every martingale is a measure free martingale, and conversely that every measure free martingale admits a finitely additive measure on a certain algebra under which it is a martingale. Conditions under which such a measure is countably additive are given. Last section is devoted to an ab initio discussion of the existence of an equivalent martingale measure and the uniqueness of such a measure if they are chosen to maximize certain relative entropies.

1.2. Relative Entropy and Gibbs-Boltzmann Measures

1.2.1. *Entropy Maximization Results*

Let $(\Omega, \mathcal{B}, \mu)$ be a measure space. A $\mathcal{B}$- measurable function $f : \Omega \to [0, \infty)$ is called a probability density function (p.d.f) with respect to μ if $\int_\Omega f(\omega)\mu(d\omega) = 1$. Then $\nu_f(A) = \int_A f(\omega)\mu(d\omega), A \in \mathcal{B}$, is a probability measure dominated by μ. The relative entropy of ν_f with respect to μ is defined as

$$H(f, \mu) = -\int_\Omega f(\omega) \log f(\omega)\mu(d\omega) \tag{1}$$

provided the integral on the right hand side is well defined, although it may possibly be infinite. In particular, if μ is a finite measure, this holds since the function $h(x) = x \log x$ is bounded on $(0,1)$ and hence $\int_\Omega(-f(\omega)\log f(\omega))^+\mu(d\omega) < \infty$. This does allow for the possibility that $H(f, \mu)$ could be $-\infty$ when $\mu(\Omega)$ is finite. We will show below that if $\mu(\Omega)$ is finite and positive then $H(f, \mu) \leq \log \mu(\Omega)$ for all p.d.f. f with respect to μ. In particular if $\mu(\Omega) = 1$, $H(f, \mu) \leq 0$.

We recall here for the benefit of the reader that a $\mathcal{B}$ measurable non-negative real valued function f always has a well defined integral with respect to μ. It is denoted by $\int_\Omega f(\omega)\mu(d\omega)$. The integral may be finite or infinite. A real valued $\mathcal{B}$ measurable function f can be written as $f = f_+ - f_-$, where, for each $\omega \in \Omega$,

$$f_+(\omega) = \max\{0, f(\omega)\}, f_-(\omega) = -\min\{0, f(\omega)\}.$$

If at least one of f_+, f_- has a finite integral, then we say that f has a well defined integral with respect to μ and write

$$\int_\Omega f(\omega)\mu(d\omega) = \int_\Omega f_+(\omega)\mu(d\omega) - \int_\Omega f_-(\omega)\mu(d\omega).$$

Now note the simple fact from calculus. The function $\phi(x) = x - 1 - \log x$ on $(0, \infty)$ has a unique minimum at $x = 1$ and $\phi(1) = 0$. Thus for all $x > 0$, $\log x \leq x - 1$ with equality holding if and only if $x = 1$. So if f_1 and f_2 are two probability density functions on $(\Omega, \mathcal{B}, \mu)$, then for all ω,

$$f_1(\omega) \log f_2(\omega) - f_1(\omega) \log f_1(\omega) \leq f_2(\omega) - f_1(\omega), \tag{2}$$

with equality holding if and only if $f_1(\omega) = f_2(\omega)$. Assume now that $f_1(\omega) \log f_1(\omega)$, $f_1(\omega) \log f_2(\omega)$ have definite integrals with respect to μ and that one of them is finite. On integrating the two sides of (2) we get

$$\begin{aligned}
&\int_\Omega f_1(\omega) \log f_2(\omega) \mu(d\omega) - \int_\Omega f_1(\omega) \log f_1(\omega) \mu(d\omega) \\
&\leq \int_\Omega f_2(\omega) \mu(d\omega) - \int_\Omega f_1(\omega) \mu(d\omega) \\
&= 1 - 1 = 0.
\end{aligned}$$

The middle inequality becomes an equality if and only if equality holds in (2) a.e. with respect to μ. We have proved:

Proposition 2.1. *Let $(\Omega, \mathcal{B}, \mu)$ be a measure space and let f_1, f_2 be two probability density functions on $(\Omega, \mathcal{B}, \mu)$. Assume that the functions $f_1 \log f_1, f_1 \log f_2$ have definite integrals with respect to μ and that one of them is finite. Then*

$$H(f_1, \mu) = -\int_\Omega f_1(\omega) \log f_1(\omega) \mu(d\omega) \leq -\int_\Omega f_1(\omega) \log f_2(\omega) \mu(d\omega), \tag{3}$$

with equality holding if and only if $f_1(\omega) = f_2(\omega)$, a.e. μ.

Note that if $\mu(\Omega)$ is finite and positive and if we set $f_2(\omega) = (\mu(\Omega))^{-1}$, for all ω, then the right hand side of (3) becomes $\log \mu(\Omega)$ and we conclude that relative entropy of $H(f_1, \mu)$ is well defined and at most $\log \mu(\Omega)$.

Let f_0 be a probability density function on $(\Omega, \mathcal{B}, \mu)$ such that $\lambda \equiv H(f_0, \mu)$ is finite and let

$$\mathcal{F}_\lambda = \{f : f \textit{ a p.d.f. wrt } \mu \textit{ and } \quad -\int_\Omega f(\omega) \log f_0(\omega) \mu(d\omega) = \lambda\}. \tag{4}$$

From Proposition 2.1 it follows that for any $f \in \mathcal{F}_\lambda$,

$$H(f,\mu) = -\int_\Omega f(\omega)\log f(\omega)\mu(d\omega) \le -\int_\Omega f(\omega)\log f_0(\omega)\mu(d\omega) = \lambda$$
$$= -\int_\Omega f_0(\omega)\log f_0(\omega)\mu(d\omega),$$

with equality holding if and only if $f = f_0$, *a.e.* μ. We summarize this as:

Theorem 2.1. *Let $f_0, \lambda, \mathcal{F}_\lambda$, be as in (4) above. Then*

$$\sup\{H(f,\mu) : f \in \mathcal{F}_\lambda\} = H(f_0,\mu)$$

and f_0 is the unique maximiser.

Theorem 2.1. says that any probability density function f_0 with respect to μ with finite entropy relative to μ appears as the unique solution to an entropy maximizing problem in an appropriate class of probability density functions. Of course, this assertion has some meaning only if $\mathcal{F}_\lambda$ does not consist of f_0 alone. A useful reformulation of this result is as follows:

Theorem 2.2. *Let $h : \Omega \to \mathbb{R}$ be a $\mathcal{B}$ measurable function. Let c and λ be real numbers such that*

$$\psi(c) = \int_\Omega e^{ch(\omega)}\mu(d\omega) < \infty, \qquad \int_\Omega \mid h(\omega) \mid e^{ch(\omega)}\mu(d\omega) < \infty, \text{ and}$$

$$\lambda = \frac{\int_\Omega h(\omega)e^{ch(\omega)}\mu(d\omega)}{\psi(c)}.$$

Let $f_0 = \frac{e^{ch}}{\psi(c)}$ and let

$$\mathcal{F}_\lambda = \{f : f \text{ a p.d.f. wrt } \mu, \text{ and } \int_\Omega f(\omega)h(\omega)\mu(d\omega) = \lambda\}.$$

Then

$$\sup\{H(f,\mu) : f \in \mathcal{F}_\lambda\} = H(f_0,\mu),$$

and f_0 is the unique maximiser.

Here are some sample examples of the above considerations.

Example 1. Let $\Omega = \{1, 2, \cdots, N\}, N < \infty$, μ counting measure on Ω, $h = 1, \lambda = 1$. Then $\mathcal{F}_\lambda = \{(p_i)_{i=1}^N, p_i \geq 0, \sum_{i=1}^N p_i = 1\}$. Then

$$f_0(j) = \frac{1}{N}, j = 1, 2, \cdots, N,$$

the uniform distribution on $\{1, 2, \cdots, N\}$, maximizes the relative entropy of the class $\mathcal{F}_\lambda$ with respect to μ.

Example 2. Let $\Omega = \mathbb{N}$, the natural numbers $\{1, 2, \cdots\}$, $\mu =$ the counting measure on Ω, $h(j) = j, j \in \mathbb{N}$. Fix $\lambda, 1 \leq \lambda < \infty$ and let

$$\mathcal{F}_\lambda = \{(p_j)_{j=1}^\infty : \forall\ j,\ p_j \geq 0, \sum_{j=1}^\infty p_j = 1, \sum_{j=1}^\infty j p_j = \lambda\}.$$

Then $f_0(j) = (1-p)p^{j-1}, j = 1, 2, \cdots$, where $p = 1 - \frac{1}{\lambda}$, maximizes the relative entropy of the class $\mathcal{F}_\lambda$ with respect to μ.

Example 3. Let $\Omega = \mathbb{R}$, $\mu =$ Lebesgue measure on $\mathbb{R}$, $h(x) = x^2$, $0 < \lambda < \infty$. Set $\mathcal{F}_\lambda = \{f : f \geq 0, \int_\mathbb{R} f(x)dx = 1, \int_\mathbb{R} x^2 f(x)dx = \lambda\}$. Then

$$f_0(x) = \frac{1}{\sqrt{2\pi\lambda}} e^{-\frac{x^2}{2\lambda}},$$

i.e., the Gaussian distribution with mean zero and variance λ, maximizes the relative entropy of the class $\mathcal{F}_\lambda$ with respect to the Lebesgue measure.

These examples are well known (see [5]) and the usual method is by the use of Lagrange's multipliers. The present method extends to the case of arbitrary number of constraints (see [1], [8]).

Definition 2.1. Let $(\Omega, \mathcal{B}, \mu)$ be a measure space. Let $h : \Omega \to \mathbb{R}$ be $\mathcal{B}$ measurable and let c be a real number. Let $\psi(c) = \int_\Omega e^{ch(\omega)} \mu(d\omega)$ be finite. Let

$$\nu_{(\mu,c,h)}(A) = \frac{\int_A e^{ch(\omega)} \mu(d\omega)}{\psi(c)}, A \in \mathcal{B}.$$

The probability measure $\nu_{(\mu,c,h)}$ is called the *Gibbs-Boltzmann measure* corresponding to (μ, c, h).

Example 4. (Spin system on N states.) Let $\Omega = \{-1, 1\}^N$, N a positive integer, and let $V : \Omega \to \mathbb{R}$, $0 < \beta_T < \infty$ be given. Let μ denote the counting measure on Ω. The measure

$$\nu_{(\mu,\beta_T,V)}(A) = \frac{\sum_{\omega \in A} e^{-\beta_T V(\omega)}}{\sum_{\omega \in \Omega} e^{-\beta_T V(\omega)}}, \qquad A \subset \Omega, \tag{5}$$

is called the *Gibbs distribution* with potential function V and temperature constant β_T for the spin system of N states. The denominator on the right side of (5) is known as the *partition function.*

1.2.2. *Weak Convergence of Gibbs-Boltzmann Distribution*

Let Ω be a Polish space, i.e., a complete separable metric space and let $\mathcal{B}$ denote its Borel σ-algebra. Recall that a sequence $(\mu_n)_{n=1}^{\infty}$ of probability measures on $(\Omega, \mathcal{B})$ converges weakly to a probability measure μ on $(\Omega, \mathcal{B})$, if

$$\int_{\Omega} f(\omega)\mu_n(d\omega) \to \int_{\Omega} f(\omega)\mu(d\omega)$$

for every continuous bounded function $f : \Omega \to \mathbb{R}$. Now let $(\mu_n)_{n=1}^{\infty}$ be a sequence of probability measures on $(\Omega, \mathcal{B})$, $(h_n)_{n=1}^{\infty}$ a sequence of $\mathcal{B}$ measurable functions from Ω to $\mathbb{R}$ and $(c_n)_{n=1}^{\infty}$ a sequence of real numbers. Assume that for each $n \geq 1$, $\int_{\Omega} e^{c_n h_n(\omega)}\mu(d\omega) < \infty$. For each $n \geq 1$, let $\nu_{(\mu_n, c_n, h_n)}$ be the Gibbs-Boltzmann measure corresponding (μ_n, c_n, h_n) as in definition 2.1. An important problem is to find conditions on $(\mu_n, h_n, c_n)_{n=1}^{\infty}$ so that $(\nu_{(\mu_n, c_n, h_n)})_{n=1}^{\infty}$ converges weakly. We address this question in a somewhat special context. We start with some preliminaries.

Let $C \subset \mathbb{R}$ be a compact subset and μ a probability measure on Borel subsets of $\mathbb{R}$ with support contained in C. For $c \in \mathbb{R}$, let

$$\psi(c) = \int_C e^{cx}\mu(dx).$$

Since C is bounded and μ is a probability measure, the function ψ is well defined and infinitely differentiable on $\mathbb{R}$. For any $k \geq 1$,

$$\psi^{(k)}(c) = \int_C e^{cx}x^k\mu(dx).$$

Note that the function $f_c(x) = \frac{e^{cx}}{\psi(c)}$ is a probability density function with respect to μ with mean $\phi(c) = \frac{\psi'(c)}{\psi(c)}$.

Proposition 2.2.

(i) *ϕ is infinitely differentiable and $\phi'(c) > 0$ for all $c \in \mathbb{R}$, provided μ is not supported on a single point. If μ is a Dirac measure, i.e., if μ is supported on a single point, then $\phi'(c) = 0$ for all c.*

(ii) $\lim_{c\to-\infty}\phi(c) = \inf\{x : \mu(-\infty, x) > 0\} \equiv a$,
(iii) $\lim_{c\to+\infty}\phi(c) = \sup\{x : \mu[x,\infty) > 0\} \equiv b$,
(iv) *for any* α, $a < \alpha < b$, *there is a unique* c *such that* $\phi(c) = \alpha$.

Proof: If μ is a Dirac measure then the claims are trivially true, so we assume that μ is not a Dirac measure. Since ψ is infinitely differentiable and $\psi(c) > 0$ for all c, ϕ is also infinitely differentiable. Moreover,

$$\phi'(c) = \frac{(\int_C x^2 e^{cx}\mu(dx))\psi(c) - (\psi'(c))^2}{(\psi(c))^2}$$

can be seen as the variance of a non-constant random variable X_c whose distribution is absolutely continuous with respect to μ with probability density function $f_c(x) = \frac{e^{cx}}{\psi(c)}$. (Note that X_c is non-constant since μ is not concentrated at a single point and f_c is positive on the support of μ.) Thus $\phi'(c) =$ variance of $X_c > 0$, for all c. This proves (i).

Although a direct verification of (ii) is possible we will give a slightly different proof. We will show that as $c \to -\infty$, the random variable X_c converges in distribution to the constant function a so that $\phi(c)$ which is the expected value of X_c converges to a. Note that by definition of a, for all $\epsilon > 0$, $\mu([a, a+\epsilon)) > 0$ while $\mu((-\infty, a)) = 0$. Also $\mu((b,\infty)) = 0$, whence

$$P(X_c > a+\epsilon) = \frac{\int_{(a+\epsilon,b]} e^{cx}\mu(dx)}{\psi(c)} = \frac{\int_{(a+\epsilon,b]} e^{c(x-a)}\mu(dx)}{\int_{[a,b]} e^{c(x-a)}\mu(dx)}.$$

For $c < 0$, and $0 < \epsilon < \frac{b-a}{2}$,

$$\begin{aligned} P(X_c > a+\epsilon) &\leq \frac{e^{c\epsilon}\mu((a+\epsilon, b])}{e^{c\frac{\epsilon}{2}}\mu([a, a+\frac{\epsilon}{2}))} \\ &= e^{c\frac{\epsilon}{2}}\frac{\mu((a+\epsilon,b])}{\mu([a,a+\frac{\epsilon}{2}))} \to 0 \quad \text{as} \quad c \to -\infty. \end{aligned}$$

Also, since $\mu((-\infty, a)) = 0$, $\quad P(X_c < a) = 0$. So, $X_c \to a$ in distribution as $c \to -\infty$, whence $\phi(c) \to a$ as $c \to -\infty$. This proves (ii). Proof of (iii) is similar. Finally (iv) follows from the intermediate value theorem since ϕ is strictly increasing and continuous with range (a, b). This completes the proof of Proposition 2.2.

Proposition 2.2 also appears at the beginning of the theory of large deviations (see [10]) thus giving a glimpse of the natural connection between large deviation theory and entropy theory. (See Varadhan's interview, p31, [11].)

The requirement that μ have compact support can be relaxed in the above proposition. The following is a result under a relaxed condition.

Let μ be a measure on $\mathbb{R}$. Let $I = \{c : \int_{\mathbb{R}} e^{cx}\mu(dx) < \infty\}$. It can be shown that I is a connected subset of $\mathbb{R}$ which can be empty, a singleton, or an interval that is half open, open, closed, finite, semi-finite or all of $\mathbb{R}$ (see [2]). Suppose I has a non-empty interior I^0. Then in I^0 the function $\psi(c) = \int_{\mathbb{R}} e^{cx}\mu(dx)$ is infinitely differentiable with $\psi^{(k)}(c) = \int_{\mathbb{R}} e^{cx}x^k\mu(dx)$. Further $\phi(c) = \frac{\psi'(c)}{\psi(c)}$ satisfies

$$\phi'(c) = \frac{\psi''(c) - (\psi'(c))^2}{(\psi(c))^2},$$

which is positive, being equal to the variance of a random variable with probability density function $\frac{e^{cx}}{\psi(c)}$ with respect to μ. Thus, for any α satisfying $\inf_{c\in I^0}\phi(c) < \alpha < \sup_{c\in I^0}\phi(c)$, there is a unique c_0 in I^0 such that $\phi(c_0) = \alpha$.

Let μ be a probability measure on $\mathbb{R}$. Note that a real number λ is the mean $\int_{\mathbb{R}} x\nu(dx)$ of a probability measure ν absolutely continuous with respect to μ if and only if $\mu(\{x : x \le \lambda\}), \mu(\{x : x \ge \lambda\})$ are both positive.

As a corollary of Proposition 2.2 we have:

Corollary 2.1. *Let the closed support of μ be a compact set C. Let α be such that $\mu\{x : x \le \alpha\}, \mu\{x : x \ge \alpha\}$ are both positive. Let*

$$\mathcal{F}_\alpha = \{f : f \text{ a pdf}, \quad \int_C xf(x)\mu(dx) = \alpha\}.$$

Then there exists a unique probability density function g with respect to μ such that

$$H(g, \mu) = \max\{H(f, \mu) : f \in \mathcal{F}_\alpha\}.$$

If $\alpha = \inf C$ or if $\alpha = \sup C$, then μ necessarily assigns positive mass to α and $g = \frac{1}{\mu(\{\alpha\})} \times 1_{\{\alpha\}}$. Let $\inf C < \alpha < \sup C$. Then there is a unique c such that with $g = f_c = \frac{e^{cx}}{\int_C e^{cx}\mu(dx)}$ one has $\alpha = \int_C xf_c(x)\mu(dx)$ and

$$H(g, \mu) = H(f_c, \mu) = \max\{H(f, \mu) : f \in \mathcal{F}_\alpha\}.$$