Electron. Commun. Probab. **31** (2026), article no. 1, 1–12.
https://doi.org/10.1214/26-ECP753
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

An analogue Fréchet-Shohat moments convergence theorem for indeterminate moment problems

Pier Luigi Novi Inverardi* Aldo Tagliani†

Abstract

The Fréchet-Shohat Theorem (FST) (1931) is a well-established result concerning determinate moment problems. We endeavor to address an analogous problem within the realm of indeterminate Hamburger and Stieltjes moment problems, focusing exclusively on absolutely continuous random variables. We demonstrate that, under an additional condition stipulating the convergence of the entropy of a sequence of distribution functions to the entropy of the unique maximum entropy distribution, a stronger mode of convergence is achieved, which subsequently implies convergence in distribution. This result is attainable due to the inherent property of indeterminate moment problems possessing a unique density g_{hmax}, distinguishable from other solutions by its maximal entropy. In conclusion, the moment convergence implies weak convergence under determinacy and under indeterminacy, if there is entropy convergence to the g_{hmax} then there is also weak convergence.

Keywords: Indeterminate Hamburger and Stieltjes moment problem; Method of Moments; Maximum Entropy; Entropy convergence; Probability distribution.
MSC2020 subject classifications: 44A60, 60E05, 62E17, 94A17, 62B10.
Submitted to ECP on June 6, 2024, final version accepted on July 17, 2025.

1 Introduction and motivation

The Fréchet-Shohat Theorem (FST) (Fréchet and Shohat, 1931, Section 6), also known as the *second limit theorem*, is a fundamental result in the literature. Expressed in terms of distribution functions (d.fs.), this theorem addresses the following question: given a sequence of d.fs. $\{G_n\}_{n\in\mathbb{N}}$, under what conditions does there exist a d.f., G, such that $\lim_{n\to\infty} G_n = G$ in distribution? Convergence in distribution signifies that $G_n(x) \to G(x)$ at every continuity point x of G. This theorem needs the satisfaction of three conditions:

(a) For all n and k, the k-th absolute moment of G_n is finite, i.e., $\int |x|^k \; dG_n < \infty$. Consequently, all moments $m_{k,n} = \int x^k \; dG_n$ are finite.

(b) As $n \to \infty$, the moments $m_{k,n}$ converge to finite values m_k, implying the existence of a d.f. G with the moment sequence $\{m_k\}$.

*University of Trento (Italy). E-mail: pierluigi.noviinverardi@unitn.it - Corresponding author
†idem. E-mail: aldo.tagliani@unitn.it

(c) The d.f. G associated with the moment sequence $\{m_k\}$ must be unique.

When conditions (a), (b), and (c) are met, the sequence G_n converges to G in distribution as $n \to \infty$. This theorem applies to sequences of d.fs. whose support is a bounded or unbounded subset of $\mathbb{R}$ or $\mathbb{R}^+$, regardless of whether the G_n are discrete, continuous, or singular.

We now pose the question: if condition (c) of the FST is relaxed, and we assume the existence of infinitely many distribution functions sharing the moment sequence $\{m_k\}$ (i.e., the moment problem is indeterminate), can the convergence in distribution of a sequence $\{G_n\}$ to a distribution G still be established? While it is generally acknowledged that convergence in distribution does not hold in the indeterminate case (as demonstrated by sequences oscillating between distinct solutions), we investigate what additional conditions imposed on the sequence of distribution functions $\{G_n\}$ would guarantee its convergence in distribution to G.

This paper provides a precise answer, leveraging the Maximum Entropy (MaxEnt) principle and recent findings by the authors on entropy convergence. Specifically, within an indeterminate moment problem, among the infinite densities consistent with the given moments, there exists a unique distribution, denoted G_{hmax}, that possesses the maximum entropy. The authors have previously demonstrated that MaxEnt approximants converge in entropy to this distribution, thus uniquely characterizing G_{hmax} by the moment sequence $\{m_k\}$. Consequently, as will be shown in Theorem 3.1, it is possible to construct sequences $\{G_n\}$ that satisfy the moment constraints and converge in entropy to G_{hmax}. Furthermore, Section 2.2 will establish that convergence in entropy implies convergence in distribution, thereby achieving our objective. Given that G_{hmax} is the sole distribution characterized by the provided moment sequence, it is the only distribution to which such sequences $\{G_n\}$ can converge, unlike other solutions that necessitate information beyond $\{m_k\}$. In this respect, the approach mirrors that of the FST, where a single distribution is characterized by its moments.

The remainder of this paper is structured as follows: Section 2 reviews pertinent results on the MaxEnt method, established theorems on convergence in entropy, and stronger modes of convergence. Section 3 presents the main result concerning the implication of moment convergence for the convergence in distribution of G_n to G, drawing an analogy to the FST for determinate moment problems. Section 4 discusses an additional crucial assumption required to ensure the convergence of the distribution when the moment problem is indeterminate.

2 Some auxiliary results

2.1 MaxEnt setup

Let X be a random variable with support $\mathcal{S}_X = \mathbb{R}$ or $\mathbb{R}^+$, distribution function G, and moment sequence $\{m_k\}$. Considering the first M moments $\{m_k\}_{k=1}^M$ of G, the maximization of the Shannon (differential) entropy $H[g] = -\int_{\mathcal{S}_X} g(x)\,\log g(x)\,dx$ yields the MaxEnt approximation f_M of the density g (Kesavan and Kapur, 1992, p. 59), given by

$$f_M(x) = \exp\left(-\sum_{k=0}^{M} \lambda_k x^k\right), \tag{2.1}$$

subject to the $(M+1)$ moment constraints

$$\int_{\mathcal{S}_X} x^k f_M(x)\,dx = m_k, \quad k=0,\dots,M, \quad \text{with} \quad m_0 = 1. \tag{2.2}$$

The differential entropy (henceforth, entropy) of f_M is then

$$H[f_M] = -\int_{\mathcal{S}_X} f_M(x)\log f_M(x)\,dx = \sum_{k=0}^{M} \lambda_k m_k. \tag{2.3}$$

For the Hamburger moment problem, the degree M of the polynomial in the exponential of the MaxEnt density f_M must be even to ensure its integrability over the real line. However, the MaxEnt formalism allows for a unified treatment of both Hamburger and Stieltjes moment problems. Consequently, to promote notational parsimony and generality, we will subsequently avoid explicit indexing of sequences related to densities, distributions, moments, or entropies, given that the Stieltjes case does not necessitate the even order restriction on M.

In essence, f_M represents the density with the maximum possible entropy among all densities sharing the first M moments $\{m_k\}_{k=1}^M$. A crucial and well-established property is that as the number of moments M increases, the uncertainty does not increase, and the sequence $\{H[f_M]\}$ is monotonically non-increasing. For both the Hamburger and Stieltjes cases, we adopt the following essential condition:

All integer moments $m_k = \int_{\mathcal{S}_X} x^k\, g(x)\,dx$, for $k = 0,1,2,\dots$, of the absolutely continuous random variable X with density g and support $\mathcal{S}_X$, exist.

Note the density g may have $-\infty$ or finite entropy $H[g]$. In the latter case it follows that the sequence $\{H[f_M]\}$ constitutes a Cauchy sequence, being monotonically non-increasing with $H[f_M] \geq H[g]$ for each M. Notably, if the density g is bounded, its entropy $H[g]$ is guaranteed to be finite. Indeed, if $0 < g(x) \leq L < \infty$ for all x, then

$$-H[g] = \int_{\mathcal{S}_X} g(x)\log g(x)\,\mathrm{d}x < \int_{\mathcal{S}_X} g(x)\log L\,\mathrm{d}x = \log L \;\Rightarrow\; H[g] \geq -\log L > -\infty.$$

Conversely, an unbounded density g does not necessarily imply infinite entropy (for instance, in Stieltjes case take $g(x) = \frac{1}{2\,r\,x^{\frac{2r-1}{2r}}} e^{-x^{\frac{1}{2r}}}$, $x > 0$, $r > 1$).
We now introduce the following classes of densities:

$$\mathcal{C}_M = \left\{ g_M \geq 0 : \int_{\mathcal{S}_X} x^k g_M(x)\,dx = m_k,\; k = 1,2,\dots,M \right\},$$

with the natural extension to

$$\mathcal{C}_\infty = \left\{ g \geq 0 : \int_{\mathcal{S}_X} x^k g(x)\,dx = m_k,\; k = 1,2,\dots \right\}.$$

It is important to note that $\mathcal{C}_M$ is a convex set for every M, and consequently, $\mathcal{C}_\infty = \bigcap_M \mathcal{C}_M$ is also a convex set. In the M-indeterminate Hamburger and Stieltjes cases, $\mathcal{C}_\infty$ contains infinitely many densities that serve as solutions to the moment problem.

Furthermore, the Shannon entropy induces an ordering principle on the elements $g \in \mathcal{C}_\infty$. Since entropy is a concave functional (Kapur (1989), p. 223), there exists a unique density with the largest entropy, which we denote as g_{hmax}, whose entropy satisfies

$$H[g_{hmax}] = \max_{g \in \mathcal{C}_\infty} H[g].$$

Considering the sequence of MaxEnt densities $\{f_M\}$ defined in (2.1), Novi Inverardi et al. (2024a) and (2024b) have demonstrated that $\{f_M\}$ converges in entropy to g_{hmax}, which is uniquely characterized by $\{m_k\}$, analogous to the determinate case. From this convergence in entropy, and through stronger modes of convergence derived from information theory, these authors ultimately proved convergence in distribution. Consequently, g_{hmax} is *entropy-distinguishable* from other members of C_∞. Moreover, given that g_{hmax} is the sole density that can be reconstructed using only the moment sequence $\{m_k\}$, it is *representative* of the entire class C_∞.

Concerning notation, in the ensuing discussion, encompassing both Hamburger and Stieltjes moment problems, we will distinguish between two categories of densities within the sets $\mathcal{C}_M$ and $\mathcal{C}_\infty$: those possessing the maximum entropy and the others. Given a moment sequence $\{m_k\}$, we establish the following definitions:

1. within $\mathcal{C}_M$: $f_M \in \mathcal{C}_M$ denotes the MaxEnt density constrained by $\{m_k\}_{k=1}^M$, with distribution function F_M (continuous for every $x \in \mathcal{S}_X$) and entropy $H[f_M]$. Conversely, $g_M \in \mathcal{C}_M$ denotes the remaining probability density functions, with distribution function G_M and entropy $H[g_M] < H[f_M]$.
2. within $\mathcal{C}_\infty$: $g_{hmax} \in \mathcal{C}_\infty$ denotes the MaxEnt density constrained by $\{m_k\}$, with distribution function G_{hmax} (continuous for each $x \in \mathcal{S}_X$) and entropy $H[g_{hmax}]$. Furthermore, $g \in \mathcal{C}_\infty$ denotes the remaining densities, with distribution function G and entropy $H[g] < H[g_{hmax}]$.

The subsequent discussion will rely fundamentally on the following two theorems, established for M-indeterminate Hamburger and Stieltjes moment problems by Novi Inverardi et al. (2024a, Thm. 3.1) and Novi Inverardi et al. (2024b, Thm. 7).

Theorem A. *In the indeterminate Hamburger or Stieltjes moment problem, the sequence $\{f_M\}$ of the MaxEnt approximations converges in entropy to g_{hmax}, i.e.,*

$$\lim_{M \to \infty} H[f_M] = H[g_{hmax}], \tag{2.4}$$

where $H[g_{hmax}]$ can be finite or $-\infty$.

Integrating Theorem A with Theorem 1 and Theorem 2 from Stoyanov et al (2024), we arrive at the following result:

Theorem B. *In the indeterminate Hamburger or Stieltjes moment problem, the infinitely many densities in $\mathcal{C}_M$ and $\mathcal{C}_\infty$ have entropies that span the two different intervals*

$$\Big(-\infty,\ H[f_M]\Big] \quad \textit{and} \quad \Big(-\infty,\ \inf_M H[f_M] = H[g_{hmax}]\Big].$$

2.2 A Hierarchy of Different Convergence Modes

Considering the class of absolutely continuous probability distributions with finite entropy, we can establish the following two preliminary results concerning MaxEnt densities f_M and g_{hmax} with corresponding distribution functions F_M and G_{hmax}. Here, we focus on the Hamburger moment problem; the Stieltjes case follows analogously,

with the exception that the requirement of "f_M having M even moments" is not imposed due to the non-negativity of the support.
Two distinct cases concerning $H[g_{hmax}]$ are separately discussed.

1. $H[g_{hmax}]$ finite.

   Given that the sequence $\{H[f_M] \geq H[g_{hmax}]\}$ is monotonically non-increasing and bounded below by the finite value $H[g_{hmax}]$, it constitutes a Cauchy sequence. Let $K > M$, and denote by f_K and f_M the maximum entropy solutions to the truncated moment problem constrained by the first K and M moments, respectively. Then:

   a) From Theorem A, the monotonically non-increasing sequence $\{H[f_M]\}$ converges to $H[g_{hmax}]$ and is therefore a Cauchy sequence.
   b) Consider the relative entropy (Kullback-Leibler divergence) between f_K and f_M, which share the same first M moments, defined as

   $$D(f_K \| f_M) = \int_{\mathcal{S}_X} f_K(x) \log \frac{f_K(x)}{f_M(x)}\, dx.$$

   Utilizing (2.1), the definition of $H[f_K]$, (2.2), and (2.3) in sequence, we obtain

   $$D(f_K \| f_M) = H[f_M] - H[f_K], \tag{2.5}$$

   where f_M is the MaxEnt density given the first M moments $\{m_k\}_{k=1}^{M}$.

   Then invoking the established relationship between the $\mathcal{L}^1$-distance of the probability density functions f_K and f_M, denoted by $\|f_K - f_M\|_1$, and the relative entropy (Kullback-Leibler divergence) $D(f_K \| f_M)$, we have the following inequality (as presented in Toussaint (1975), equation (5)):

   $$\frac{1}{2}\|f_M - f_K\|_1^2 \;\leq\; D(f_K \| f_M) \;=\; H[f_M] - H[f_K]. \tag{2.6}$$

   Given Theorem A and the assumption that the density g_{hmax} has finite entropy, the sequence $\{f_M\}$ is a Cauchy sequence in $\mathcal{L}^1$ due to (2.6). Consequently, taking into account (2.6), we have

   $$\sup_x |F_M(x) - F_K(x)| \leq \|f_M - f_K\|_1 \leq \sqrt{2(H[f_M] - H[f_K])}, \tag{2.7}$$

   which implies that $\{F_M\}$ is also a Cauchy sequence with respect to the uniform norm. Replacing f_K with g_{hmax} and F_K with G_{hmax} in (2.7), and recalling Theorem A, we obtain

   $$\lim_{M\to\infty} \sup_x |F_M(x) - G_{hmax}(x)| = 0, \tag{2.8}$$

   demonstrating that $\{F_M\}$ converges uniformly to G_{hmax}. Hence, $\{F_M\}$ also converges pointwise to G_{hmax}:

   $$\lim_{M\to\infty} F_M(x) = G_{hmax}(x) \tag{2.9}$$

   for every $x \in \mathcal{S}_X$.

2. $H[g_{hmax}] = -\infty$.

   From Toussaint (1975), eqn.(6), proceeding similarly to (2.5) - (2.7), replacing f_K with g_{hmax} and taking into account Theorem A, we have

   $$|F_M(x) - G_{hmax}(x)| \leq \|f_M - g_{hmax}\|_1 \leq \sqrt{2(D(g_{hmax}\|f_M))} = \sqrt{2(H[f_M] - H[g_{hmax}])}, \tag{2.10}$$

   so that $\{F_M\}$ converges pointwise to G_{hmax}, in analogy with (2.9).

3 Main result

We now present the central result of this paper, applicable to both the Hamburger and Stieltjes indeterminate moment problems. This is accomplished through a refined set of initial assumptions, concerning the sequence of distributions $\{G_M\} \to G_{hmax}$. This reinforcement is required by the presence of infinitely many densities within the set $\mathcal{C}_\infty$, only one of which is uniquely identifiable through its entropy. Consequently, given solely the assigned moments, the sequence $\{G_M\}$ is constrained to converge to G_{hmax} as its unique limit.

Theorem 3.1 (Novi Inverardi and Tagliani)**.** *Let $\{m_k\}$ be an indeterminate moment sequence and let $\{g_M\}$ be probability densities satisfying*

$$\int x^k g_M(x)\, dx = m_k, \quad k = 0, 1, \ldots, M. \tag{3.1}$$

If $M \to \infty$, $H[g_M] \to H[g_{hmax}]$, where $H[g_{hmax}]$ can be finite or $-\infty$, then the sequence $\{G_M\}$ converges uniformly (or, pointwise) in distribution to G_{hmax}.

Note that the k-th moment of G_M is m_k when $k \leq M$, so the moment convergence condition is satisfied.

Proof.

1. $H[g_{hmax}]$ finite

   Note that the construction of densities $g_M \in \mathcal{C}_M$ that generate the required sequence $\{G_M\}$ with a specific finite entropy value, thereby satisfying condition 2 of Theorem 3.1, is detailed in the proof of Theorem 1 in Stoyanov et al. (2024). The crux of the proof lies in the construction of specific scenarios concerning a sequence of distribution functions $\{G_M\}$ that satisfy to the two requisite conditions stipulated by the theorem.

   - **Scenario 1**: This scenario considers sequences $\{g_M\}$ wherein each element g_M possesses entropy $H[g_M]$ satisfying the inequality $H[g_{hmax}] \leq H[g_M] < H[f_M]$, a viable assumption based upon Theorem B. Theorem A establishes that $\lim_{M\to\infty} H[f_M] = H[g_{hmax}]$, from which it logically follows that $\lim_{M\to\infty} H[g_M] = H[g_{hmax}]$. The desired sequence $\{G_M\}$, fulfilling both conditions of the theorem, is then obtained through a constructive methodology detailed in Stoyanov et al (2024), in the proof of Theorem 1. Alternatively, one may assume the choice $\{g_M = f_M\}$, which implies $\{H[g_M] = H[f_M]\}$. In this instance, the required sequence $\{G_M = F_M\}$ is derived from the fact that $\lim_{M\to\infty} H[f_M] = H[g_{hmax}]$. It is noteworthy that in both selections of $\{G_M\}$, the second assumption of the theorem is inherently satisfied due to Theorem A and thus does not require explicit consideration. Moreover, in both procedures, the sequence $\{G_M\}$ is obtained constructively.

   - **Scenario 2**: This scenario assumes a sequence $\{g_M\}$ that satisfies the following two conditions:

     i) The entropy $H[g_M]$ of g_M lies within the interval $\big(-\infty, H[g_{hmax}]\big]$.

     ii) $\lim_{M\to\infty} H[g_M] = H[g_{hmax}]$.

In contrast to Scenario 1., the second condition must be explicitly imposed. The feasibility of both conditions is guaranteed by Theorems A and B. Analogous to Scenario 1, the sequence $\{G_M\}$ is obtained through a constructive procedure.

In conclusion, the validity of the above scenarios is conditional on the existence of a single density $g_{hmax} \in \mathcal{C}_\infty$ characterised by a finite maximum entropy $H[g_{hmax}]$ and distinguishable, in terms of entropy, from all other densities $g \in \mathcal{C}_\infty$.

Having established these preliminaries, we now proceed to the proof of Theorem (3.1), employing the sequences $\{g_M\}$ constructed in the preceding two scenarios. Consider the relative entropy $D(g_M \| f_M)$ between f_M and g_M, where both densities share the first M moments $\{m_k\}_{k=1}^M$, and f_M represents the maximum entropy density. Invoking equation (2.1), the definition of $H[g_M]$, equation (2.2), and equation (2.3) in succession, we obtain the relation $D(g_M \| f_M) = H[f_M] - H[g_M]$. Subsequently, utilizing the well-established inequality relating relative entropy to the $\mathcal{L}^1$-distance between f_M and g_M (see Toussaint (1975), eq. (5)), we have:

$$\frac{1}{2} \parallel f_M - g_M \parallel_1^2 \leq D(g_M \| f_M) = H[f_M] - H[g_M],$$

from which the following inequalities are derived:

$$\sup_x \mid F_M(x) - G_M(x) \mid \leq \parallel f_M - g_M \parallel_1 \leq \sqrt{2D(g_M \| f_M)} = \sqrt{2(H[f_M] - H[g_M])}.$$

By substituting G_M with G_{hmax} and g_M with g_{hmax} respectively, recalling Theorem A, the assumption $H[g_M] \to H[g_{hmax}]$, as $M \to \infty$, for each considered sequence $\{g_M\}$, the completeness of the $\mathcal{L}^1$ space, and finally taking the limit as $M \to \infty$, the proof of Theorem 3.1 is thus concluded.

2. $H[g_{hmax}] = -\infty$

   In this case there is only one scenario and the proof of Theorem 3.1 comes from (2.10) taking into account Theorem A.

□

In light of the preceding observation concerning the FST, condition $\lim_{M\to\infty} H[g_M] = H[g_{hmax}]$ of Theorem 3.1 is assumed to be merely sufficient for the theorem's validity. Indeed, the possibility of orderings of the components of C_∞, beyond that dictated by entropy, opens the possibility for suitable modes of convergence that, in turn, yield convergence in distribution.

4 Discussion about condition $\lim_{M\to\infty} H[g_M] = H[g_{hmax}]$ of Theorem 3.1

The FST, valid for the determinate case, relies exclusively on convergence in entropy of Theorem 3.1 spent in its formalization. Conversely, the indeterminate case is characterized by the existence of an infinite number of admissible distributions. A subset of these distributions is typically constructed through the application of Stieltjes classes technique (for definition and properties, refer to Stoyanov (2004)). This construction starts assuming the existence of a distribution G possessing the prescribed moments $\{m_k\}$, followed by the determination of a perturbation function h that leaves the moments of the combination $g^*(x) = g(x) + \varepsilon \cdot g(x) \cdot h(x)$, $\varepsilon \in [-1, 1]$, invariant in the sense that $\int x^k dG^*(x) = m_k$ for all $k \geq 0$. This procedure generates an infinite family of distributions G^*, the Stieltjes class indeed, belonging to $\mathcal{C}_\infty$.

Within the framework of indeterminate moment problems, we would like to highlight two specific contributions that merit attention and closer examination.

a) The foundational work of Nevanlinna (1922) introduced the use of specific matrices to characterize the entire set of solutions. This framework, known as the Nevanlinna parametrization (see Berg (1995), for a comprehensive review), establishes a bijective mapping between the class of all measures satisfying a prescribed sequence of moments and the class of Pick functions (also referred to as Nevanlinna or Herglotz functions), denoted by φ. The practical application of this parametrization relies on four specific entire functions $A(z), B(z), C(z)$, and $D(z)$, collectively known as the Nevanlinna functions. These functions are uniquely determined by the underlying indeterminate moment sequence. Detailed treatments of these functions can be found in the seminal monographs by Akhiezer (1965) and Shohat and Tamarkin (1943), where they are identified as fundamental tools for investigating the properties of indeterminate measures. Despite their theoretical importance, the number of cases where the functions A, B, C, and D have been explicitly calculated remains limited. The first explicit examples were derived in the 1990s through the contributions of Ismail, Masson, Berg, Valent, Chihara, and Christiansen, among others. A contemporary overview of these cases and their associated references is provided in Schmüdgen (2017), Section 7.

b) More recently, focusing specifically on density functions, Novi Inverardi and Tagliani (2024a, 2024c) and Novi Inverardi et al. (2024b) proposed an ordering principle for the elements of $\mathcal{C}_\infty$ by exploiting the concavity of the Shannon entropy functional. They established the existence of a unique density, g_{hmax}, which possesses a strictly higher entropy than any other distribution in $\mathcal{C}_\infty$. Within the Nevanlinna parametrization, identifying the specific Pick function φ that corresponds to g_{hmax} remains an open and compelling problem. Furthermore, the authors demonstrated the convergence in entropy of MaxEnt approximants to this maximal density, that is $\lim_{M\to\infty} H[f_M] = H[g_{hmax}]$, thereby uniquely characterizing g_{hmax} by the given moment sequence $\{m_k\}$. It follows that, within the MaxEnt framework - particularly for the purpose of distribution reconstruction - determinate and indeterminate moment problems may be viewed as equivalent, as the reconstruction process in both scenarios converges to a single, well-defined distribution. Consequently, in an indeterminate setting where the available information is restricted to the moment sequence, g_{hmax} is the only distribution that can be uniquely recovered. The identification of any other element $g \in \mathcal{C}_\infty$ would necessitate auxiliary information, such as the specification of an appropriate perturbation function h within a Stieltjes class construction, as discussed above.

With the following simple example we prove that the convergence of moments implies convergence in distribution.

Example 1

Consider within Scenario 1 the sequence $\{g_n \equiv f_n\}$ in which the condition

$$\lim_{n\to\infty} H[g_n] = H[g_{hmax}]$$

of Theorem 3.1 is verified by Theorem A. The convergence in distribution comes from Theorem 3.1 and consequently FST holds also in the M-indet case.

We present a further example that is instructive since it clearly highlights the role played by condition $\lim_{M\to\infty} H[f_M] = H[g_{hmax}]$ of Theorem 3.1, concerning convergence in entropy. More precisely, we now aim to illustrate, through a further example having general validity, that the omission of convergence of entropy in above mentioned theorem implies there is no convergence in distribution. Specifically, any sequence that alternates between two or more distinct solutions to the moment problem, will necessarily fail to converge.

Example 2

Let X be a r.v. on $\mathbb{R}^+$ with standard LogN distribution, so $X \sim LogN$, g be its density with d.f. G and G_n, $n = 1, 2, ...$ a sequence of d.fs., with density g_n and $H[g_{hmax}]$ finite.

a) The following example, due to Prof. Jordan Stoyanov, allows us to examine what happens when the requirement of convergence in entropy - a key part of Theorem 3.1 - is removed. Assume

$$g_n(x) = g(x)[1 + \frac{1}{2}(-1)^n \sin(2\pi \log x)], \quad x > 0$$

For any n, the r.v. $X_n \sim G_n$ has finite moments $m_{k,n} \to m_k$, the moments of X. It is known that LogN, with entropy $H[LogN] = \frac{1}{2}\log(2\pi e)$, is M-indet. Look at the sequence $\{G_n\}$. We find the following

$$G_{2n} \to G(x) + \frac{1}{2}\Delta(x) \quad \text{and} \quad G_{2n-1} \to G(x) - \frac{1}{2}\Delta(x)$$

where $\Delta(x) = \int_0^x g(u)\sin(2\pi \log u)\, du$. Then the two sub-sequences converge to different limits. Therefore, $\{G_n\}$ does not converge.

We now examine the aforementioned example from the perspective of entropy, being the only criterion known to us. Indeed, being condition $\lim_{n\to\infty} H[g_n] = H[g_{hmax}]$ only sufficient, it makes no sense to evaluate the previous example in entropic terms. But, having shown that $\{g_n\}$ does not converge, it is sufficient to conclude that such a condition on $\{g_n\}$ must be imposed. Initially, we observe that the moments employed constitute the Lognormal moment sequence. In this instance, Novi Inverardi and Tagliani (2024c) - Theorem 2 have demonstrated that the density g coincides with the density of the Lognormal distribution that is, $LogN = g_{hmax}$. Both subsequences $H[g_{2n}]$ and $H[g_{2n-1}]$ have constant entropies for each $n \in \mathbb{N}$, as well such entropies are slight different, $\mid H[g_{2n}] - H[g_{2n-1}] \mid \sim 10^{-8}$. Both $H[g_{2n}]$ and $H[g_{2n-1}$ have value $\simeq 1.3543$, which is strictly less than the entropy of g_{hmax}, $H[g_{hmax}] = \frac{1}{2}\log(2\pi e) \simeq 1.4189$. Consequently, the sequence $\{g_n\}$ does not converge in entropy to g_{hmax}. Indeed, formulating $g_n(x)$ as $g_n(x) = g(x) + \frac{1}{2}(-1)^n g(x)\sin(2\pi \log x)$, the sequence $\{g_n\}$ includes information coming from the given moment sequence through the terms $g(x)$ and an additional one, independent on n, through the term $\frac{1}{2}(-1)^n g(x)\sin(2\pi \log x)$.

Drawing upon the implications of Theorem 3.1, we can assert that the aforementioned results are a consequence of omitting the condition $\lim_{M\to\infty} H[f_M] = H[g_{hmax}]$.

b) We now proceed to prove the applicability of Theorem 3.1 under scenario 2, employing the conditions of convergence in entropy stipulated therein.

Consider a sequence $\{G_n\}$ constructed to align with the specifications of scenario 2 in Theorem 3.1, while maintaining an invariant sequence of moments. This is achieved through a subtle perturbation of the underlying probability density functions as follows:

$$g_n(x) = g_{hmax}(x) \cdot [1 + \frac{1}{2}(-a)^n \sin(2\pi \log(x))], \quad x > 0, \ \mid a \mid < 1 \tag{4.1}$$

so that $H[g_{hmax}]$ is finite. Examining the sequence $\{G_n\}$, we observe that, in analogy to the preceding item a), as $n \to \infty$, the difference $\Delta(x)$ approaches zero:

$$\Delta(x) = a^n \int_0^x g_{hmax}(u) \sin(2\pi \log(u))\, du \to 0.$$

Consequently, both subsequences $\{G_{2n}\}$ and $\{G_{2n-1}\}$ converge to G.

It remains to establish the validity of entropy convergence condition. The elements g_n belong to a Stieltjes class centred at g_{hmax}, with a perturbation function $h(x) = \frac{1}{2}(-a)^n \sin(2\pi \log(x))$, such that we can express $g_n(x) = g_{hmax}(x)[1 + h(x)]$. The function $h(x)$ represents additional information that vanishes as $n \to \infty$, according with condition ii) of Scenario 2.

Given that g_n and g_{hmax} share the same sequence of moments, it follows that:

$$\int_{\mathbb{R}^+} x^k g_{hmax}(x) h(x)\ dx = 0, \quad k \geq 0 \tag{4.2}$$

Consider now the Kullback-Leibler distance

$$D(g_n \| g_{hmax}) = \int_{\mathbb{R}^+} g_n(x) \log \frac{g_n(x)}{g_{hmax}(x)}\ dx = \int_{\mathbb{R}^+} g_{hmax}(x)[1 + h(x)] \log[1 + h(x)]\ dx.$$

Recalling the elementary inequality $\log(x) \leq x - 1$ for positive real numbers, which implies $x \log(x) \leq x^2 - x$, and considering equation (4.2) with $k = 0$, we find that as $n \to \infty$:

$$\begin{aligned} D(g_n \| g_{hmax}) &\leq \int_{\mathbb{R}^+} g_{hmax}(x)[h^2(x) + h(x)]\, dx \\ &= \int_{\mathbb{R}^+} g_{hmax}(x)[h^2(x)]\, dx \leq \frac{1}{4} a^{2n} \int_{\mathbb{R}^+} g_{hmax}(x)\ dx = \frac{1}{4} a^{2n} \to 0 \end{aligned}$$

Thus, the sequence $\{g_n\}$ converges to g_{hmax} in relative entropy that is,

$$\lim_{n \to \infty} D(g_n \| g_{hmax}) = 0. \tag{4.3}$$

Considering that g_n and f_M belong to $\mathcal{C}_M$, and that f_M is the MaxEnt density, we have $D(g_n \| f_M) = H[f_M] - H[g_n]$. By Theorem A (not provided here) and equation (2.7), it is implied that $\{H[f_M]\}$ and $\{f_M\}$ are both Cauchy sequences. Consequently, f_M converges almost everywhere (a.e.) to g_{hmax}. Therefore, from the convergence in relative entropy established above, as $n \to \infty$:

$$D(g_n \| g_{hmax}) = \lim_{M \to \infty} D(g_n \| f_M) = \lim_{M \to \infty} H[f_M] - H[g_n] = H[g_{hmax}] - H[g_n] \to 0.$$

As a direct consequence of the assumed form of g_n in equation (4.1), we have:

$$\lim_{n\to\infty} H[g_n] = H[g_{hmax}].$$

And this demonstrates that under the condition $\lim_{M\to\infty} H[f_M] = H[g_{hmax}]$ of Theorem 3.1, the sequence $\{G_n\}$ converges in entropy, and consequently in distribution uniformly to $G \equiv G_{hmax}$, in accordance with scenario 2.

In conclusion,

a) when relying solely on the information derivable from the moment sequence, the issues pertaining to determinate and indeterminate cases prove equivalent. In both scenarios, indeed, one arrives at the reconstruction of a single distribution. This distribution coincides with g in the determinate case or g_{hmax}, both uniquely characterized by the given moment sequence $\{m_k\}$ and representing the non-committal distribution, in accordance with Jaynes' principle that requires "...*we should not assume more than what we know*".

b) The infinite multitude of further distributions belonging to $\mathcal{C}_\infty$ in the indeterminate case can be obtained solely by integrating supplementary information beyond that conveyed by the moments. Taking all of this into account, it is not unexpected that an analogue to the FST may be formulated for indeterminate moment problems as well.

References

[1] Akhiezer, N.I., *The Classical Moment Problem and Some Related Questions in Analysis*, Oliver and Boyd, Edinburgh, (1965).

[2] Berg, C., "Indeterminate moment problems and the theory of entire functions". *J. Comput. Appl. Math.*, **65**, no. 1-3, 27–55, (1995).

[3] Fréchet M., Shohat J., "A Proof of the Generalized Second-Limit Theorem in the Theory of Probability". *Transactions of the American Mathematical Society*, **33**(2), 533-543, (1931).

[4] Jaynes, E.T. , "Information theory and statistical mechanics". *Phys. Rev.*, **106**, 620, (1957).

[5] Kapur J.N., *Maximum-entropy Models in Science and Engineering*, Wiley Eastern Limited, (1989)

[6] Kesavan H.K., Kapur J.N., *Entropy Optimization Principles with Applications*, Academic Press, London, (1992)

[7] Nevanlinna R., "Asymptotische Entwicklungen beschrankter Funktionen und das Stieltjessche Momentenproblem". *Ann. Acad. Sci. Fenn. A*, **18**, 1-53, (1922).

[8] Novi Inverardi P.L., Tagliani A., "Indeterminate Stieltjes moment problem: entropy convergence". *Symmetry*, **16**, 313, (2024a). https://doi.org/10.3390/sym16030313

[9] Novi Inverardi P.L., Tagliani A., Milev, M., "Indeterminate Hamburger moment problem: entropy convergence". *Stat. Prob. Lett.*, **212**, 110155, (2024b). https://doi.org/10.1016/j.spl.2024.110155

[10] Novi Inverardi P.L., Tagliani A., "The Lognormal distribution is characterized by its integer moments" *Mathematics*, **12**(23), 3830, (2024c). https://doi.org/10.3390/math12233830

[11] Schmüdgen, K., *The Moment Problem*. Graduate Texts in Mathematics **277**, Springer, Cham, (2017).

[12] Shohat, J.A., Tamarkin, J.D., *The Problem of Moments*, Math. Surveys No. **1**. Amer. Math. Soc.: Providence (RI), (1943).

[13] Stoyanov J., "Stieltjes Classes for Moment-Indeterminate Probability Distributions", *J. of Appl. Probab.*, **41A**, 281–294, (2004).

[14] Stoyanov J.M., Tagliani A., Novi Inverardi P.L., "Maximum Entropy Criterion for Moment Indeterminacy of Probability Densities". *Entropy*, *26*, 121, (2024).

[15] Toussaint, G.T., "Sharper Lower Bounds for Discrimination Information in Terms of Variation". *IEEE Transactions on Information Theory*, **21** (1), 99-100, January (1975).

Acknowledgments. The authors are grateful to the Associate Editor and all anonymous reviewers for their thoughtful suggestions, which significantly improved the final exposition. Special thanks are due to Professor Jordan Stoyanov for providing a compelling example that clarified the relevance of the convergence in entropy condition, an insight that proved essential to the development of Theorem 3.1.