When Do the Moments Uniquely Identify a Distribution

Carlos A. Coelho, Rui P. Alberto, and Luís M. Grilo

Abstract The authors establish when do the moments $E(X^h)$, for h in some subset C of $\mathbb{R}$, uniquely identify the distribution of any positive random variable X, that is, when is x^h a separating function. The simple necessary and sufficient condition is shown to be related with the existence of the moment generating function of the random variable $Y = log X$. The subset C of $\mathbb{R}$ is thus the set of values of h for which the moment generating function of Y is defined. Examples of random variables characterized in this way by the set of their h-th moments are given.

Keywords Moment problem · Identifiability · Separating functions · Log-normal distribution · F distribution

1 Introduction

The possibility of identifying a distribution from the expression of its moments is of unquestionable importance in statistics. There are areas where this is indeed an extremely important and useful tool, being used for decades, as it is the case of the study and characterization of distributions of l.r.t. (likelihood ratio test) statistics

---

Funded by Portuguese national funds through FCT–Fundação para a Ciência e a Tecnologia, I.P., through NOVA Math, under the scope of projects UIDB/00297/2020 and UIDP/00297/2020 and through CIMA, under project UIDP/04674/2020.

C. A. Coelho (✉)
Mathematics Department and NOVA Math, NOVA School of Science and Technology, NOVA University of Lisbon, Caparica, Portugal
e-mail: cmac@fct.unl.pt

R. P. Alberto · L. M. Grilo
NOVA Math, NOVA School of Science and Technology, NOVA University of Lisbon, Caparica, Portugal

L. M. Grilo
Mathematics Department and CIMA (Research Center in Mathematics and Applications), Évora University, Évora, Portugal

S. Kumar et al. (eds.), *Directional and Multivariate Statistics*,
https://doi.org/10.1007/978-981-96-2004-3_13

in Multivariate Analysis. Very often one is able to obtain first the expressions for the moments of such statistics, from which the distribution is then inferred (see for example Kshirsagar, 1972, Chap. 8; Coelho & Marques, 2010; Coelho & Arnold, 2019, Chap. 5). In this setup the work is done under a very precise framework, which is the fact that we are dealing with r.v.'s (random variables) with support or range in $(0, 1)$. It is a well-known fact that r.v.'s with a delimited support or range have their distributions completely determined by their moments, more precisely, by the set of their positive integer order moments, a fact that is clearly and simply stated by Anant Kshirsagar in his landmark book in Multivariate Analysis, when he says "In general, the moments of a distribution do not determine the distribution uniquely. … However, under certain conditions, the moments do determine a distribution uniquely. A finite range is a sufficient condition for this purpose." (Kshirsagar, 1972, Chap. 8, Sect. 2), a fact that derives directly from the so-called Hausdorff moment problem (Hausdorff, 1921a, 1921b, 1923).

The Hausdroff moment problem is the problem of finding the necessary and sufficient conditions a given sequence of positive values $\{m_n\}$ $(n \in \mathbb{N}_0)$ has to verify to be taken as the sequence of moments of a r.v. X with support or range $(0, 1)$. The answer is that such a sequence has to verify the relation $m_0 = 1$, and (Akhiezer, 1965, Chap. 2, Sect. 6.4; Feller, 1971, Chap. VII, Sect. 3)

$$\sum_{i=0}^{k}(-1)^i \binom{k}{i} m_{n+i} \geq 0\,,$$

for all $k, n \in \mathbb{N}_0$. If the sequence $\{m_n\}$ verifies such condition, we have $m_n = E(X^n)$, for $n \in \mathbb{N}_0$, and the distribution of X is the only one that has such integer order moments. The result may be extended to r.v.'s with any delimited support or range (Kshirsagar, 1972, Chap. 8, Sect. 2; Schmüdgen, 2017, Theorems 3.13, 3.14). In practice, Hausdorff's result is most commonly used in the following way: once obtained an expression for the positive integer order moments of a r.v. with a delimited support or range, which yields well-defined finite values for all integer order moments, or once asserted that a r.v. with such a support has all positive integer order moments defined and finite, we may then infer that the distribution of this r.v. is uniquely determined by this set of moments and eventually deduce the distribution of the r.v. (see for example Khatri, 1965; Kshirsagar, 1972, Chap. 8, Sect. 2; Coelho & Arnold, 2019, Chap. 5). Usually it is said that the Hausdorff moment problem is a well-determined moment problem, since if such a sequence may be identified as a sequence of positive integer order moments of a r.v. with delimited support, then such a r.v. is unique, or equivalently, the distribution that yields such moments is unique.

But, there are other two moment problems, which, in common language, may be either determined or undetermined. These are the Hamburger moment problem (Hamburger, 1920a, 1920b, 1921; Widder, 1946, Chap. III, Sect. 10; Shoat & Tamarkin, 1970, Chap. I, II) and the Stieltjes moment problem (Stieltjes, 1894, 1895; Widder, 1946, Chap. III, Sect. 13; Shoat & Tamarkin, 1970, Chap. I). The Hamburger moment problem refers to r.v.'s with support or range $\mathbb{R}$, and the Stieltjes moment problem to r.v.'s with support or range $\mathbb{R}^+$ or $\mathbb{R}_0^+$. In very simple terms, in both cases the question is: "given a sequence $\{m_n\}$ ($n \in \mathbb{N}_0$), may it be taken as the set of positive integer order moments of some r.v., or rather, of some distribution?". In case of an affirmative answer to this question, the following question that arises is: "is that r.v. or distribution unique?". If the answer to both questions is affirmative the moment problem is said to be determinate or determined and if the answer to the first question is affirmative but the answer to the second question is negative, the problem is said to be indeterminate or undetermined.

We know that if a r.v. X has a m.g.f. (moment generating function), $M_X(t) = E(e^{tX})$, then all of its positive integer order moments are finite and well-defined, and the distribution of this r.v. is uniquely identified by the set of these moments. In this case the positive integer order moments, say $\mu_n = E(X^n)$, for $n \in \mathbb{N}_0$, yield, for t in some neighborhood of zero, the relation

$$M_X(t) = \sum_{n=0}^{\infty} \frac{t^n \mu_n}{n!} \quad \text{for } t \in V_{\epsilon,\epsilon'}(0), \tag{1}$$

where $V_{\epsilon,\epsilon'}(0) = \{x : x \in \left] -\epsilon, \epsilon' \right[; \epsilon, \epsilon' > 0\}$ represents a neighborhood of zero. That is, the set of positive order moments uniquely identifies a distribution whenever the series on the right hand side of (1) is convergent for t in a neighborhood of zero.

We should however note that, on one hand, the existence of the m.g.f. for a given r.v. implies that its c.f. (characteristic function) is analytic at $t = 0$, while, on the other hand, the lack of compliance with the convergence of the series in (1) does not imply that the r.v. is not uniquely determined by the set of its positive integer order moments.

For r.v.'s that although not having a m.g.f. still have moments of all positive integer orders, thus having an analytic c.f. at $t = 0$, we may think about checking for the applicability of the Carleman condition (Carleman, 1926; Akhiezer, 1965, Chap. 2, Probl. 11) or Carleman type conditions (Akhiezer, 1965, Chap. 2; Lin, 1997; Wu, 2002). See Appendix A for the Carleman condition and a, by far not exhaustive but rather limited, list of Carleman type conditions.

But it still remains to consider two important types of r.v.'s. And it is to these r.v.'s that the result in the present work is dedicated. These are r.v.'s with support in $\mathbb{R}^+$ or $\mathbb{R}_0^+$ (or, in fact, in $\mathbb{R}^-$ or $\mathbb{R}_0^-$), which

1. do not have positive integer order moments of all orders, or even of no positive integer order, or
2. have integer order moments of all orders but these fully match the integer order moments of other r.v.'s.

The question is if anyway the distributions of these r.v.'s are still fully determined by the set of available moments.

We may note that in the definition of $E\left(X^h\right)$, which, using the notation of the Stieltjes integral, is given by

$$E\left(X^h\right) = \int_S x^h dF_X(x),$$

for all cases where $\int_S |x|^h dF_X(x)$ is convergent, where S is the support of the r.v. X and $F_X(x)$ its c.d.f. (cumulative distribution function), nothing goes against to consider $h \in \mathbb{R}$ or h in some subset of $\mathbb{R}$.

Let us consider a r.v. in the situation 1 above. Let X be a r.v. with an $F_{m,n}$ distribution. We then have

$$E(X^h) = \frac{\Gamma\left(\frac{m}{2}+h\right)\,\Gamma\left(\frac{n}{2}-h\right)}{\Gamma\left(\frac{m}{2}\right)\,\Gamma\left(\frac{n}{2}\right)}\left(\frac{n}{m}\right)^h, \tag{2}$$

which is valid for $-m/2 < h < n/2$. The question is: does this set of moments uniquely identify the distribution of X?

A sequence of r.v.'s in the situation 2 above is brought to our attention by Stieltjes himself (Stieltjes, 1894, 1895, Chap. VIII). Let us consider the sequence of r.v.'s X_n, ($n \in \mathbb{N}_0$), with p.d.f.'s (probability density functions)

$$f_{X_n}(x) = \frac{e^{-\frac{1}{2}(\log x)^2}}{\sqrt{2\pi}x}\{1+\alpha\sin(2n\pi\log x)\}, \quad |\alpha| \le 1 \tag{3}$$

which for $n = 0$ and/or $\alpha = 0$ yields the well known standard Log-Normal distribution. For any $h \in \mathbb{R}$ we have

$$\begin{aligned} E\left(X_n^h\right) &= \int_0^\infty x^h \frac{e^{-\frac{1}{2}(\log x)^2}}{\sqrt{2\pi}x}\{1+\alpha\sin(2n\pi\log x)\}dx \\ &= e^{h^2/2}\underbrace{\int_0^\infty \frac{e^{-\frac{1}{2}(\log x - h)^2}}{\sqrt{2\pi}x}\{1+\alpha\sin(2n\pi\log x)\}dx}_{=1+e^{-\frac{1}{2}(2n\pi)^2}\alpha\sin(2n\pi h)} \\ &= e^{h^2/2}\left\{1+e^{-\frac{1}{2}(2n\pi)^2}\alpha\sin(2n\pi h)\right\}, \end{aligned} \tag{4}$$

where the details in the computation of the integral were left aside since they are a bit lengthy while not being the key point in this brief note. For $h \in \mathbb{N}_0$ expression (4) yields

$$E\left(X_n^h\right) = e^{h^2/2}, \quad \forall h \in \mathbb{N}_0, \tag{5}$$

which is neither a function of n nor of α (see for example Feller, 1971, Chap. VII, Sect. 3; Knight, 2000, Sect. 1.7 and also Casella & Berger, 2002, Sect. 2.3, for the cases $n = 0$ and $n = 1$), so that all r.v.'s X_n ($n = 0, 1, 2, \ldots$) have exactly the same set of positive integer order moments. But the p.d.f.'s in (3) are quite different for different values of n and α.

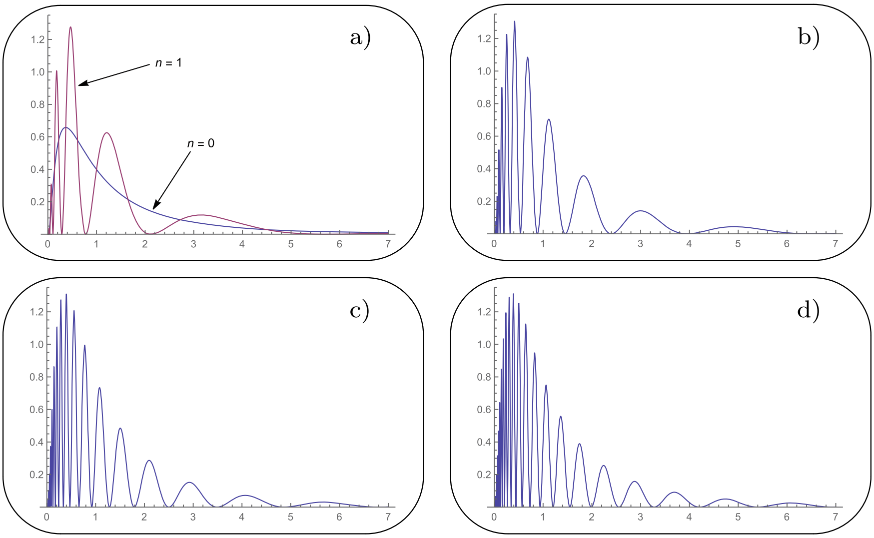

Fig. 1 Plots of the p.d.f.'s in (3) for $\alpha = 1$ and consecutive values of n: a) $n = 0$ and $n = 1$; b) $n = 2$; c) $n = 3$; d) $n = 4$. All plots have the same horizontal and vertical scales

The p.d.f.'s in (3) yield very different plots for different values of n and α, as it may be seen in Figs. 1 and 2.

So, the question that remains is exactly similar to the one asked above: does the whole set of moments of each X_n uniquely identify the distribution of each r.v. X_n?

So, actually the question to be asked is: is it possible to devise a result or results that can be used to assert that the set of available moments for these r.v.'s uniquely identifies their distributions? And then the further question that may bear in our minds

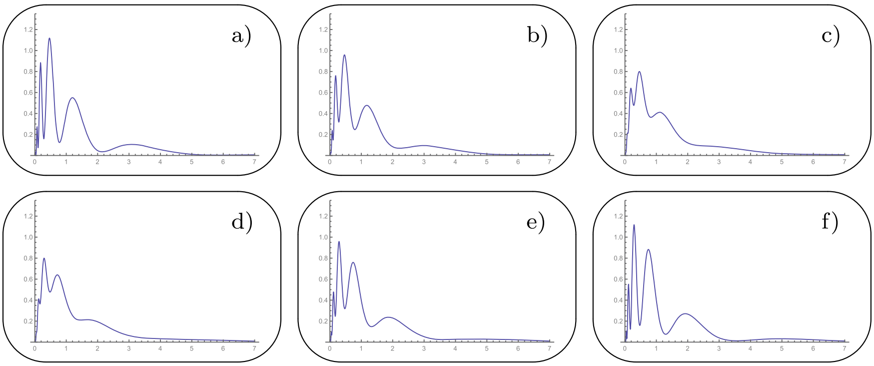

Fig. 2 Plots of the p.d.f.'s in (3) for $n = 1$ and different values of α: a) $\alpha = 0.75$; b) $\alpha = 0.5$; c) $\alpha = 0.25$; d) $\alpha = -0.25$; e) $\alpha = -0.5$; f) $\alpha = -0.75$. All plots have the same horizontal and vertical scales (see Fig. 1 for the plot corresponding to $n = 1$ and $\alpha = 1$)

may be: and is it that the same result may be used for both cases 1 and 2 mentioned above, although they may seem at first to correspond to quite different cases?

The answers to these questions are given in the next Section.

2 The Main Results

In what concerns expression (4) above, we may easily see that $E\left(X_n^h\right)$ is indeed a function of both n and α, as long as $2nh \notin \mathbb{Z}$ (i.e., $2nh \in \mathbb{R}\backslash\mathbb{Z}$), which is usually an overlooked detail. Indeed, in this case we may actually say that although for $h \in \mathbb{Z}$, $E\left(X_n^h\right)$ is given by (5), thus being neither a function of n nor of α, for $h \in \mathbb{R}$ the moments $E\left(X_n^h\right)$ uniquely identify the distributions in (3).

In simple terms, the function $w(\cdot)$ is said to be separating if and only if for any two c.d.f.'s $F(\cdot)$ and $G(\cdot)$,

$$\int_{\mathbb{R}} w(x)dF(x) = \int_{\mathbb{R}} w(x)dG(x) \Longrightarrow F(x) = G(x), \forall x \in \mathbb{R}.$$

First of all let us state that, for $i = \sqrt{-1}$, and $h \in \mathbb{R}$, the functions x^{ih} are always separating for the distributions of r.v.'s with support $\mathbb{R}^+$, since if the r.v. X has support $\mathbb{R}^+$we may always define the r.v. $Y = \log X$, with

$$E\left(X^{ih}\right) = E\left(e^{ih\log X}\right) = E\left(e^{ihY}\right) = \Phi_Y(h), \quad h \in \mathbb{R}, \tag{6}$$

which shows that the moments $E\left(X^{ih}\right)$, with $h \in \mathbb{R}$, will always exist and will uniquely identify the distribution of X, provided that X has support $\mathbb{R}^+$, since then there will be a r.v. $Y = \log X$, which will have c.f. $\Phi_Y(h) = E\left(X^{ih}\right)$. As it happens for the c.f. (and actually also with the m.g.f.), one usually is not much interested in individual values but rather in the expression of the c.f. or of the m.g.f. itself, since by looking at it, or by using an inverse transform, it is possible to identify the distribution of X and/or to obtain the expression of its p.d.f. or p.m.f. (probability mass function), or of its c.d.f.. $E\left(X^{ih}\right)$ is called the Mellin transform of X (Mellin, 1887, 1899, 1900, 1902, 1904, 1910; Bertrand et al., 2010; Debnath & Bhatta, 2015, Chap. 8; Coelho & Arnold, 2019, Chap. 2, Sect. 2.1).

Then, using a similar argument and since the m.g.f. of Y, when it exists, also uniquely identifies the distribution of Y, we may say that if $M_Y(h) = E\left(e^{hY}\right)$ exists, similarly to (6),

$$M_Y(t) = E\left(e^{tY}\right) = E\left(e^{t\log X}\right) = E\left(X^t\right), \quad t \in V_{\epsilon,\epsilon'}(0)$$

and vice-versa, if X is a r.v. with support in $\mathbb{R}^+$, with $E(X^h)$ defined for h in a neighborhood of zero, $V_{\epsilon,\epsilon'}(0)$, then the r.v. $Y = \log X$ is defined and it has a m.g.f. for $t \in V_{\epsilon,\epsilon'}(0)$, with

$$V_{\epsilon,\epsilon'}(0) = \left\{ x \in \mathbb{R} : x \in]-\epsilon, \epsilon'[\, ; \epsilon, \epsilon' \in \overline{\mathbb{R}}^+ \right\}$$

where $\overline{\mathbb{R}}^+ = \mathbb{R}^+ \cup \{+\infty\}$, with $V_{-\infty,\infty}(0) = \mathbb{R}$. All we have to do is to look at $E(X^h)$ as a function of h and if $E(X^h)$ is defined for h in a neighborhood of zero, then it uniquely identifies the distribution of X, or, in other words, in that case, X is the only r.v. with that set of moments.

We may thus first establish the following Lemma, which states that, if for a given r.v. $X > 0$, $E\left(X^h\right)$ exists (and is finite) for some $h > 0$, then $E(X^r)$ exists for any $r \in]0, h]$ and if $E\left(X^h\right)$ exists (and is finite) for some $h < 0$ then $E(X^r)$ also exists for any $r \in [h, 0[$.

Lemma 1 *If X is a non-negative random variable such that $E(X^{h_1}) < \infty$ for some $h_1 \in (-\infty, 0)$ and $E(X^{h_2}) < \infty$ for some $h_2 \in (0, \infty)$, then $E(X^r) < \infty$, $\forall r \in (h_1, h_2)$.*

Proof If $r \in (0, h_2)$ then we have $X^r < 1 + X^{h_2}$ so that $E(X^r) < \infty$, while, if $r \in (h_1, 0)$ we have $(1/X)^{|r|} < 1 + (1/X)^{|h_1|}$, and we can conclude, in this case, that $E(X^r) = E((1/X)^{|r|})$ is finite since $E(X^{h_1}) = E((1/X)^{|h_1|})$ is finite. □

Then, we may state the following Proposition.

Proposition 1 *If X is a r.v. with support in $\mathbb{R}^+$ and $E\left(X^h\right)$ is defined for some $h_1 < 0$ and some $h_2 > 0$ then $E\left(X^h\right)$ is defined for $h \in [h_1, h_2]$ and, taken for the whole range of values of h for which $E\left(X^h\right)$ is defined, will uniquely determine the distribution of X.*

Proof From Lemma 1 we have that $E(X^h)$ is defined for $h \in [h_1, h_2]$, so that if we take $Y = \log X$, its m.g.f. will be given by

$$M_Y(h) = E\left[e^{hY}\right] = E\left[e^{h \log X}\right] = E\left[X^h\right], \quad h \in V_{\epsilon,\epsilon'}(0),$$

where $[h_1, h_2] \subseteq V_{\epsilon,\epsilon'}(0)$, so that in this case the r.v. $Y = \log X$ has for sure a m.g.f., being as such uniquely characterized by $E\left[e^{hY}\right]$ and as such also the distribution of $X = e^Y$, being uniquely defined as a function of Y, given the fact that the exponential and the logarithm are both one-to-one functions, is also uniquely characterized by $E\left[e^{hY}\right] = E\left[X^h\right]$, for $h \in V_{\epsilon,\epsilon'}(0)$, where $V_{\epsilon,\epsilon'}(0)$ represents the whole set of values of h for which $E(X^h)$ is defined. □

Some examples of application of the result in Proposition 1 follow.

Example 1 Going back to our first example in the previous section, let us consider $X \sim F_{m,n}$. Then, as noted, $E(X^h)$ is given by (2), for $-m/2 < h < n/2$, and as such by using the result in Proposition 1, the set of moments $E(X^h)$, for $-m/2 < h < n/2$, uniquely identifies the distribution of X. Or, in other words, the distribution $F_{m,n}$ is the only distribution with such moments, for given $m, n \in \mathbb{N}$ (actually for any $m, n \in \mathbb{R}^+$).

Example 2 Going back to our second example in the previous section, let us consider the r.v.'s X_n with the generalized Log-normal distributions with p.d.f.'s given by (3). Then we will have $E(X_n^h)$ given by (4), for $h \in \mathbb{R}$. As such, using the result in Proposition 1, we may say that the moments $E(X_n^h)$ uniquely identify the distribution of each X_n, for a given $n \in \mathbb{N}$ and a given α (with $|\alpha| < 1$). Indeed in this case the r.v.'s $Z_n = \log X_n$ have what we may call 'generalized' Normal distributions with p.d.f.'s given by

$$f_{Z_n}(z) = \frac{e^{-z^2/2}}{\sqrt{2\pi}}(1 + \alpha \sin(2n\pi z)) \tag{7}$$

and m.g.f.'s

$$\begin{aligned} M_{Z_n}(t) &= E\left(e^{tZ_n}\right) = E\left(X_n^t\right) \\ &= e^{t^2/2}\left(1 + e^{-\frac{1}{2}(2n\pi)^2}\alpha \sin(2n\pi t)\right), \quad t \in \mathbb{R} \end{aligned} \tag{8}$$

so that the functions x^h with $h \in \mathbb{R}$ are separating for the distributions of the r.v.'s X_n and the moments $E\left(X_n^h\right)$, for $h \in \mathbb{R}$, uniquely identify the distributions of the r.v.'s X_n, for any $n \in \mathbb{N}_0$ and any given α, with $|\alpha| \leq 1$. Some authors note that the m.g.f.'s in (8) yield very close numerical values for different values of n (McCullagh, 1994; Waller, 1995), although the p.d.f.'s in (7) have many different shapes (see Fig. 3). Anyway, as functions of t, the m.g.f.'s in (8) are able to fully separate the p.d.f.'s in (7). See Appendix B for a brief note on some numerical issues related with the m.g.f.'s in (8).

Example 3 The folded Cauchy distribution. Let X have a standard Cauchy distribution and let $Y = |X|$. Then Y has what we call a folded Cauchy distribution, with p.d.f.

$$f_Y(y) = \frac{2}{\pi(1 + y^2)}, \quad y > 0,$$

and

$$E(Y^h) = \sec\left(\frac{\pi h}{2}\right), \quad -1 < h < 1,$$

which shows in the light of Proposition 1 that also in this case $E(Y^h)$ (for $-1 < h < 1$) uniquely identify the distribution of Y. We should note that Y, similar to X, does not even have an expected value, but this is no problem in using the result in Proposition 1, as already observed in Example 1.

Example 4 The folded Student T distribution. Let X have a T_n distribution, and let $Y = |X|$. Then Y has what we call a folded T distribution, with n degrees of freedom, with p.d.f.

$$f_Y(y) = \frac{2}{B\left(\frac{n}{2}, \frac{1}{2}\right)} n^{-\frac{1}{2}} \left(1 + \frac{y^2}{n}\right)^{-\frac{n+1}{2}}, \quad y > 0,$$

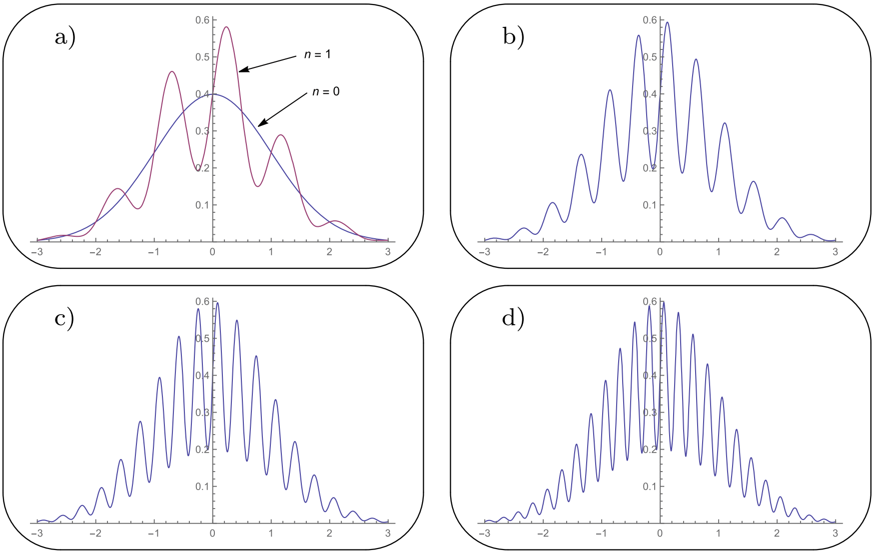

Fig. 3 Plots of the p.d.f.'s in (7) for consecutive values of n: a) $n = 0$ and $n = 1$; b) $n = 2$; c) $n = 3$; d) $n = 4$. All plots have the same horizontal and vertical scales

and

$$E(Y^n) = \frac{n^{h/2}\,\Gamma\left(\frac{h+1}{2}\right)\,\Gamma\left(\frac{n-h}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{n}{2}\right)}, \qquad -1 < h < n,$$

which shows, once again, that using Proposition 1 we may assert that the moments $E(Y^h)$ (for $-1 < h < n$) uniquely identify the distribution of Y. We may observe that for $n = 1$, X has a standard Cauchy distribution and Y a folded Cauchy distribution.

Example 5 Let X have a $Beta(a, b)$ distribution. Then

$$E(X^h) = \frac{\Gamma(a+h)\,\Gamma(a+b)}{\Gamma(a)\,\Gamma(a+b+h)}, \quad h > -a.$$

It is a well-known fact that, using Hausdorff's result, the distribution of X is fully determined by the set of its positive integer order moments, which existence is easy to spot from the expression for $E(X^h)$ above. Anyway, we may also use in this case the result in Proposition 1 to assert that, since the expression for $E(X^h)$ is valid for h in a neighborhood of zero, $E(X^h)$ uniquely identifies the distribution of X.

Example 6 Let X be a r.v. with support in $\mathbb{R}^+$, with p.d.f.

$$f_X(x) = \frac{\lambda\, e^{-\lambda \frac{1}{x}}}{x^2}, \quad \lambda > 0;\; x > 0.$$

We may note that $Y = 1/X$ has an Exponential distribution with parameter λ. Then we have

$$E(X^h) = \lambda^h \Gamma(1-h), \quad h < 1,$$

and as such the distribution of X is fully determined by the moments above.

Example 7 Let X have a standard Pareto distribution, that is, let X be a r.v. with support $(1, +\infty)$, with p.d.f.

$$f_X(x) = \frac{\alpha}{x^{\alpha+1}}, \quad \alpha > 0; \ x > 1.$$

Then

$$E(X^h) = \frac{\alpha}{\alpha - h}, \quad h < \alpha,$$

so that, from Proposition 1 we may say that the above set of moments uniquely identifies the distribution of X.

3 Conclusions

As a conclusion we may say that the answer to the question'When are the functions x^h $(h \in C \subset \mathbb{R})$ separating or, equivalently, when do the moments $E\left(X^h\right)$ $(h \in C \subset \mathbb{R})$ uniquely identify the distribution of X and how is the set C defined?' is:

(i) whenever X has support $\mathbb{R}^+$, it is thus possible to define the r.v. $Y = \log X$ (or $Y = -\log X$), and,
(ii) the moments $E(X^h)$ are defined for $h \in C \subset \mathbb{R}$, where $C = V_{\epsilon,\epsilon'}(0)$, that is, the moments $E(X^h)$ are defined for h in some neighborhood of zero, since in this case the r.v. $Y = \log X$ will have a m.g.f. $M_Y(h) = E\left(X^h\right)$ for $h \in C = \left]-\epsilon, \epsilon'\right[= V_{\epsilon,\epsilon'}(0)$, $\left(\epsilon, \epsilon' > 0\right)$;

so that we may state that

- if X has support $\mathbb{R}^+$ and we take $\epsilon < 0$ to be the smallest value of h for which $E\left(X^h\right)$ is defined, and $\epsilon' > 0$ as the largest value of h for which $E\left(X^h\right)$ is defined, then $E\left(X^h\right)$, for $h \in [\epsilon, \epsilon']$, determines the distribution of X; or we may even say that it defines the distribution of X, since then we will be able to use a Laplace inversion, at least numerically (Abate & Valkó, 2004; Cohen, 2007), to obtain the distribution of $Y = \log X$, from which the distribution of X will be readily at hand,

and thus we may say that

- x^h is separating for the distributions of r.v.'s in $\mathbb{R}^+$, if it is defined for $h \in V_{\epsilon,\epsilon'}(0)$, where $V_{\epsilon,\epsilon'}(0)$ is the set of all values h for which $E\left(X^h\right)$ is defined (that is, the set of all values h for which $\int_S \left|x^h\right| dF_X(x)$ converges), or we may also just say that

- if X_1 and X_2 are two r.v.'s with support $\mathbb{R}^+$such that $E\left(X_1^h\right) = E\left(X_2^h\right)$ for every $h \in V_{\epsilon,\epsilon'}(0)$, where $V_{\epsilon,\epsilon'}(0)$ is the set of all values h for which $E\left(X_1^h\right)$ is defined, then the r.v.'s X_1 and X_2 have the same distribution, the reason being that in this case the r.v.'s $Y_1 = \log X_1$ and $Y_2 = \log X_2$ will have the same f.g.m., and as such will have the same distribution.

We may say that if we consider the real order moments, of both positive and negative orders, these moments characterize the distribution they come from in a more insightful way than the positive integer order moments. This fact is indeed related to the fact that, as for example Ortigueira et al. (2005) and Machado (2003) state, "integer-order derivatives depend only on the local behavior of a function, while fractional order derivatives depend on the whole history of the function". If we are concerned with the fact that it may seem not possible to obtain such real order moments from the c.f. or the m.g.f. as we usually do with the integer order moments, we should note that we only have to consider the real order derivatives of those functions in order to obtain the real order moments, and that such real order derivatives if taken as the extension for $h \in \mathbb{R}$ of

$$\left.\frac{\partial^h}{\partial t^h}\Phi_X(t)\right|_{t=0} = i^h \int_S x^h dF_X(x),$$

will not only agree with the usual Grünwald-Letnikov definition of non-integer order derivative (Samko et al., 1993, Chap. 4, § 20; Mathai & Haubold, 2017, Sect. 6.4) as well as with other more general definitions as the Cauchy convolutional definition of derivative in Ortigueira et al. (2005).

Appendix 1

A Brief Note on the Carleman Condition and Other Carleman Type Conditions

The aim of this Appendix is to briefly refer to the Carleman condition and some of the Carleman type conditions available in the literature. For the proofs we refer the reader to the references cited.

The Carleman condition (Carleman, 1926; Akhiezer, 1965, Chap. 2, Problem 11) is usually referred as the result in the following Theorem.

Theorem 1 *Let $\mu_r = E\left(X^r\right) < \infty$. If*

$$\sum_{r=1}^{\infty} \frac{1}{\mu_{2r}^{1/2r}} = +\infty$$

then the distribution of the r.v. X is uniquely determined by its positive integer order moments. ■

We should note that the result in (1), although of the same type, of the result in Theorem 1, is indeed stronger, since

Another condition is the one presented in Theorem 2 (Shiryaev, 1996, Chap. II, Sect. 12, Theorem 7).

Theorem 2 *Let $\mu_r^* = E(|X|^r) < \infty$. If*

$$\limsup_{r\to\infty} \frac{(\mu_r^*)^{1/r}}{r} < \infty,$$

then the distribution of X is determined by its positive integer order moments. ■

We should note that the condition in the hypothesis of Theorem 2 implies indeed that the c.f. of the r.v. X has to be analytic at the origin. Feller (1971, Chap. XV, Sect. 4, Appendix) states a similar result, although using $\mu_r = E(X^r)$ instead of μ_r^*, and states in case that relation is verified, the c.f. of X "is analytic in a neighborhood of any point of the real axis, and hence completely determined by its power series about the origin", which is equivalent to say that the distribution of X is completely determined by the sequence of its positive order moments.

Wu (2002) presents a pair of simple but useful results, the first of them derived from Theorem 1, by applying the D'Alembert criterium.

Theorem 3 *Let X be a r.v. with $E(X^r) < \infty, \forall r \in \mathbb{N}$. If*

$$\lim_{r\to\infty} \frac{1}{r}\left|\frac{\mu_{r+1}}{\mu_r}\right| < \infty$$

then the distribution of the r.v. X is uniquely determined by its positive integer order moments. ■

And for discrete r.v.'s, derived from Theorem 2 above we have the following result.

Theorem 4 *Let X be a discrete r.v. with support N and*

$$p_k = P(X = k) \quad \forall k \in N$$

(and as such with $\sum_{k=1}^{\infty} p_k = 1$). If there is $\alpha \geq 1$ such that $p_k = O\left(e^{-k^\alpha}\right)$, that is, such that $\forall k \in N$, p_k/e^{-k^α} remains bounded, then the distribution of the r.v. X is uniquely determined by the set of its positive integer order moments. ■

In a slightly different framework, Lin (1997) obtains four criteria, the first two of which refer to r.v.'s with support $\mathbb{R}$, while the last two are for r.v.'s with support $\mathbb{R}^+$. Two of these conditions refer to r.v.'s with a differentiable p.d.f.

Theorem 5 *Let X be a r.v. with an absolutely continuous c.d.f. and p.d.f. $f(x) > 0, \forall x \in \mathbb{R}$. If $E(X^r) < \infty, \forall r \in \mathbb{N}$ and*

$$\int_{-\infty}^{\infty} \frac{-\log f(x)}{1+x^2} dx < \infty \tag{9}$$

then the distribution of the r.v. X is not uniquely determined by its positive integer order moments. ∎

Theorem 6 *Let X be a r.v. with an absolutely continuous c.d.f. and p.d.f. $f(x) > 0, \forall x \in \mathbb{R}$, symmetrical about zero and differentiable in $\mathbb{R}$, such that $E(X^r) < \infty, \forall r \in \mathbb{N}$, with*

$$f(x) \underset{x\to\infty}{\longrightarrow} 0, \quad -x\frac{f'(x)}{f(x)} \underset{x\to\infty}{\longrightarrow} \infty$$

and

$$\int_{-\infty}^{\infty} \frac{-\log f(x)}{1+x^2} dx = \infty$$

then the distribution of the r.v. X is uniquely determined by its positive integer order moments. ∎

As pointed out by Lin (1997), a sufficient condition for (9) to hold is the logarithmic mean function

$$g(t) = \frac{1}{2t}\int_{-t}^{t} |\log f(x)| dx \quad (t \in \mathbb{R})$$

to be bounded in $\mathbb{R}$.

Theorem 5 bears a close relation with Theorems I and II in Paley and Wiener (1934). Lin (1997) refers that while it was already proved by Akhiezer (1965, p. 87), using a result from Krein (1945), he proves it using the Hardy space theory.

Also as Lin (1997) shows in the proof of Theorem 6, its hypothesis is a necessary condition for the Carleman condition to be satisfied.

Using the above two theorems it is possible to prove that if X is a normally distributed r.v., its distribution is uniquely determined by its positive integer order moments, while the distributions of $X^{2n+1} (n \in \mathbb{N})$ are not (Berg, 1988).

For r.v.'s in $\mathbb{R}^+$we have the following two theorems.

Theorem 7 *Let X be a r.v. with an absolutely continuous c.d.f. and p.d.f. $f(x) > 0, \forall x \in \mathbb{R}^+$ and $f(x) = 0, \forall x \in \mathbb{R}_0^-$. If $E(X^r) < \infty, \forall r \in \mathbb{N}$ and*

$$\int_{0}^{\infty} \frac{-\log f(x^2)}{1+x^2} dx < \infty$$

then the distribution of the r.v. X is not uniquely determined by its positive integer order moments. ■

Theorem 8 *Let X be a r.v. with an absolutely continuous c.d.f. and p.d.f. $f(x) > 0, \forall x \in \mathbb{R}^+$ and $f(x) = 0, \forall x \in \mathbb{R}_0^-$, differentiable in $\mathbb{R}^+$, such that $E(X^r) < \infty, \forall r \in \mathbb{N}$, with*

$$f(x) \underset{x\to\infty}{\longrightarrow} 0, \quad -x\frac{f'(x)}{f(x)} \underset{x\to\infty}{\longrightarrow} \infty$$

and

$$\int_0^\infty \frac{-\log f\left(x^2\right)}{1+x^2} dx = \infty$$

then the distribution of the r.v. X is uniquely determined by its positive integer order moments. ■

Stoyanov et al. (2020) recently devised some new conditions for moment determinacy of r.v.'s with p.d.f.'s with support either in $\mathbb{R}$ or $\mathbb{R}^+$. We state their Theorems 1–4 in Theorems 9 through 12, with slight adaptations in the writing, to better fit in the wording of the previous theorems in this Appendix. The two first theorems are for continuous r.v.'s, and the last two for discrete r.v.'s.

Theorem 9 *Suppose the density of X is symmetric on $\mathbb{R}$ and continuous and strictly positive outside an interval $(-x_0, x_0)$, $x_0 > 1$, such that the following condition holds:*

$$K_*[f] = \int_{|x|\geq x_0} \frac{-\log\, f(x)}{x^2\, \log |x|}\, dx = \infty.$$

Further, let f be such that

$$\frac{-\log\, f(x)}{\log x} \nearrow \infty \;\; as \;\; x_0 \leq x \to \infty.$$

Then the moments of X satisfy Carleman's condition, and hence the distribution of X is fully determined by the positive integer order moments of X. Moreover, the same happens with X^2. ■

Theorem 10 *Assume that the density g of X is continuous and strictly positive on $[a, \infty)$ for some $a > 1$ such that the following condition holds:*

$$K_*[g] = \int_a^\infty \frac{-\log\, g(x^2)}{x^2\, \log x}\, dx = \infty.$$

In addition, let g be such that

$$\frac{-\log\ g(x)}{\log x} \nearrow \infty \ \ as \ \ a \leq x \to \infty.$$

Then the moments of X satisfy Carleman's condition, and hence the distribution of X is fully determined by the positive integer order moments of X. ■

Theorem 11 *Suppose that the discrete r.v. X has a p.m.f.* $f_X(x) = P(X = x)$*, with support* $\mathbb{Z}$ *that is symmetric about the origin, and that all its [positive integer order] moments are finite, and the following condition holds:*

$$\sum_{|j| \geq j_0} \frac{-\log\ f_X(j)}{j^2\ \log |j|} = \infty.$$

Here $j_0 \geq 2$*, and we assume further that*

$$\frac{-\log\ f_X(j)}{\log\ j} \nearrow \infty \ \ as \ \ j_0 \leq j \to \infty.$$

Then the moments of X satisfy Carleman's condition, and hence the distribution of X is fully determined by the positive integer order moments of X. Moreover, the same happens with X^2*.* ■

Theorem 12 *Suppose that the discrete r.v. X, with p.m.f.* $f_X(x) = P(X = x)$*, with support* $\mathbb{N}_0$*, and that all its [positive integer order] moments are finite, and the following condition holds:*

$$\sum_{n \geq n_0} \frac{-\log\ f_X(n)}{n^2\ \log n} = \infty.$$

Here $n_0 \geq 2$*, and we assume further that*

$$\frac{-\log(f_X(n)/2)}{\log\ n} \nearrow \infty \ \ as \ \ n_0 \leq n \to \infty.$$

Then the moments of X satisfy Carleman's condition, and hence the distribution of X is fully determined by the positive integer order moments of X. Moreover, the same happens with X^2*.* ■

Appendix 2

A Brief Note on Some Numerical Issues

McCullagh (1994) states that the m.g.f.'s of the r.v.'s X_n in (8) are numerically virtually indistinguishable (for $\alpha = 1/2$ and $n = 0$, $n = 1$) and Waller (1995), states that

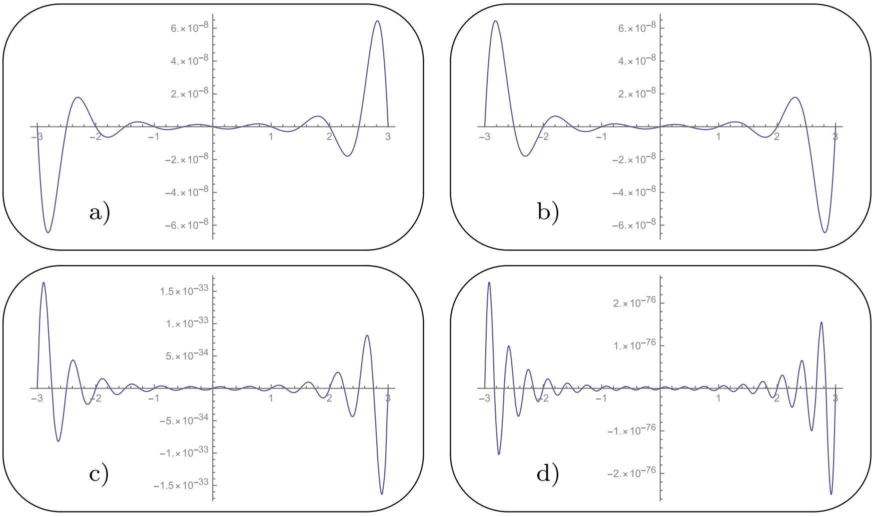

Fig. 4 Plots of the differences of m.g.f.'s in (8), of the distributions with p.d.f.'s in (7), for $-3 \leq t \leq 3$ and for consecutive values of n: a) $n = 0$ and $n = 1$; b) $n = 1$ and $n = 2$; c) $n = 2$ and $n = 3$; d) $n = 3$ and $n = 4$. All plots have the same horizontal scale but different vertical scales

only the imaginary part of the c.f. helps in this numerical discrimination. However, there is a clear scale problem when plotting the m.g.f.'s in (8). If instead we plot the difference between the values of the two m.g.f.'s. and we take the plot far enough in terms of the absolute value of t it is no longer true that the two m.g.f.'s look the same (see Fig. 4). Actually, for example for $\alpha = 1/2$, $t = 9.9$ the two functions differ by something like 1.507163×10^{12} while for $t = 12.7$ they differ by something like 1.343462×10^{26}.

We should note that the above are also exactly the differences between the moments of order $h = 9.9$ and $h = 12.7$ for the r.v.'s X_0 and X_1 with p.d.f.'s in (3) for $\alpha = 1/2$ while if we take $\alpha = 1$ the corresponding differences will be roughly the double.

We should be aware that if we do not have enough precision in the computation process we may not be able to spot such differences since although being quite large in magnitude they are quite small when compared with the original values.

Anyway, we should consider that although the numerical values of such functions may be quite close, they are analytically different for different values of n.

References

Abate, J., & Valkó, P. P. (2004). Multi-precision Laplace transform inversion. *International Journal for Numerical Methods in Engineering, 60*, 979–993

Akhiezer, N. I. (1965). *The classical moment problem and some related questions in analysis*. Translated by N. Kemmer.
Berg, C. (1988). The cube of a Normal distribution is indeterminate. *The Annals of Probability, 16*, 910–913.
Bertrand, J., Bertrand, P., & Ovarlez, J.-P. (2010). Mellin transform. In A. D. Poularikas (Eds.), *Transforms and applications handbook*, Chap. 12, 3rd edn. CRC Press, Taylor & Francis.
Billingsley, P. (2012). *Probability and measure* (Anniversary). J. Wiley & Sons Inc.
Carleman, T. (1926). *Les Fonctions Quasi-Analytiques*. Collection Borel.
Casella, G., & Berger, R. L. (2002). *Statistical inference* (2nd ed.). Duxbury.
Coelho, C. A., & Arnold, B. C. (2019). *Finite forms representations for Meijer G and Fox H functions—applied to multivariate likelihood ratio tests using Mathematica®, Maxima and R. Lecture notes in statistics*. Springer.
Coelho, C. A., & Marques, F. J. (2010). Near-exact distributions for the independence and sphericity likelihood ratio test statistics. *Journal of Multivariate Analysis, 101*, 583–593.
Cohen, A. M. (2007). *Numerical methods for Laplace transform inversion. Numerical Methods and Algorithms*, Vol. 5. Springer.
Debnath, L., & Bhatta, D. (2015). *Integral transforms and their applications* (3rd ed.). CRC Press.
Feller, W. (1971). *An introduction to probability theory and its applications*, Vol. II, 2nd edn. J. Wiley & Sons, Inc.
Hamburger, H. (1920). Bemerkungen zu einer Fragestellung des Herrn Pólya. *Mathematische Zeitschrift, 7*, 302–322.
Hamburger, H. (1920). Uber eine Erweiterung der Sieltjesschen Momentproblems. *Mathematische Annalen, 81*, 235–319.
Hamburger, H. (1921). Uber die Riemannsche Funktionalgleichung der ζ-Funktion. *Mathematische Zeitschrift, 10*, 240–254.
Hausdorff, F. (1921). Summationsmethoden und Momentfolgen. I. *Mathematische Zeitschrift, 9*, 74–109.
Hausdorff, F. (1921). Summationsmethoden und Momentfolgen. II. *Mathematische Zeitschrift, 9*, 280–299.
Hausdorff, F. (1923). Momentprobleme für ein endliches Intervall. *Mathematische Zeitschrift, 16*, 220–248.
Khatri, C. G. (1965). Classical statistical analysis based on a certain multivariate complex gaussian distribution. *The Annals of Mathematical Statistics, 36*, 98–114.
Knight, K. (2000). *Mathematical statistics*. Chapman & Hall/ CRC.
Krein, M. (1945). On a Problem of Extrapolation of A. N. Kolmogoroff. *Comptes Rendus de l'Académie des Sciences de l'URSS, XLVI, 8*, 306–309.
Kshirsagar, A. (1972). *Multivariate analysis*. Dekker.
Lin, G. D. (1997). On the moment problems. *Statistics and Probability Letters, 35*, 85–90.
Machado, J. A. T. (2003). A probabilistic interpretation of the fractional-order differentiation. *Fractional Calculus & Applied Analysis, 6*, 73–80.
Mathai, A. M., & Haubold, H. J. (2017). *An introduction to fractional calculus*. Nova Science Publishers.
McCullagh, P. (1994). Does the moment-generating function characterize a distribution? *The American Statistician, 48*, 208.
Mellin, H. (1887). Über einen zusammenhang zwischen gewissen linearen differential- und differenzengleichungen. *Acta Mathematica, 9*, 137–166.
Mellin, H. (1899). Über eine Verallgemeinerung der Riemannscher Funktion $\zeta(s)$. *Acta Societatis Scientiarum Fennica, 24*, 1–50.
Mellin, H. (1900). Eine Formel für den Logarithmus transcendenter Functionen von endlichem Geschlecht. *Acta Societatis Scientiarum Fennica, 29*, 1–49.
Mellin, H. (1902). Eine Formel für den Logarithmus transcendenter Functionen von endlichem Geschlecht. *Acta Mathematica, 25*, 165–183.

Mellin, H. (1904). Die Dirichlettschen Reihen, die zahlentheoretischen Funktionen die unendlichen Produkte von endlichem Geschlecht. *Acta Mathematica, 28*, 37–64.

Mellin, H. (1910). Abriß einer einheitlichen Theorie der Gamma- und hypergeometrischen Funktionen. *Mathematische Annalen, 68*, 305–337.

Ortigueira, M. D., Tenreiro Machado, J. A., & Sá da Costa, J. (2005). Which differintegration? *IEE Proceedings—Vision, Image, and Signal Processing, 152*, 846–850.

Paley, R. E. A. C., & Wiener, N. (1933). Notes on the theory and applications of Fourier transforms I-II. *Transactions of the American Mathematical Society, 35*, 348–355.

Samko, S. G., Kilbas, A. A., & Marichev, O. I. (1993). *Fractional integrals and derivatives—theory and applications*. Gordon and Breach Science Publishers.

Schmüdgen, K. (2017). *The moment problem. Graduate texts in mathematics*. Springer.

Shiryaev, A. N. (1996). *Probability*. Springer-Verlag.

Shoat, J. A., & Tamarkin, J. D. (1970). *The problem of moments, 4th printing*. American Mathematical Soc.

Stieltjes, T. J. (1894, 1895). Recherches sur les Fractions Continues. *Annales de la Faculté des Sciences de Toulouse, 8*, 1–122; *Annales de la Faculté des Sciences de Toulouse, 9*, A, 1–47. In W. Kapteyn & J. C. Kluyver (Eds.), *Oeuvres Complètes de Thomas Jan Stieltjes, Tome II* (pp. 402–566). Published by P. Noordhoff, Groningen, 1918.

Stoyanov, J. M., Lin, G. D., & Kopanov, P. (2020). New checkable conditions for moment determinacy of probability distributions. *Theory of Probability & Its Applications, 65*, 497–509.

Waller, L. A. (1995). Does the characteristic function numerically distinguish distributions? *The American Statistician, 49*, 150–152.

Widder, D. V. (1946). *The Laplace Transform, 2nd printing*. Princeton Univ. Press.

Wu, C. L. (2002). On the moment problem. M.Sc Thesis, Department of Applied Mathematics, National Sun Yat-Sen University, Kaohsiung, Taiwan.