Characterizations of continuous univariate probability distributions with applications to goodness-of-fit testing

S. Betsch and B. Ebner

October 16, 2018

Abstract

By extrapolating the explicit formula of the zero-bias distribution occurring in the context of Stein’s method, we construct characterization identities for a large class of absolutely continuous univariate distributions. Instead of trying to derive characterizing distributional transformations that inherit certain structures for the use in further theoretic endeavours, we focus on explicit representations given through a formula for the distribution function. The results we establish with this ambition feature immediate applications in the area of goodness-of-fit testing. We draw up a blueprint for the construction of tests of fit that include procedures for many distributions for which little (if any) practicable tests are known.

1 Introduction.

Over the last decades, Stein’s method for distributional approximation has become a viable tool for proving limit theorems and establishing convergence rates. At it’s heart lies the well-known Stein characterization which states that a real-valued random variable $Z$ has a standard normal distribution if, and only if,

$$E[f'(Z) - Zf(Z)] = 0$$ \hspace{1cm} (1.1)
holds for all functions \( f \) of a sufficiently large class of test functions. To exploit this characterization for testing the hypothesis

\[ H_0 : \mathbb{P}^X \in \{ \mathcal{N}(\mu, \sigma^2) \mid (\mu, \sigma^2) \in \mathbb{R} \times (0, \infty) \} \tag{1.2} \]

of normality, where \( \mathbb{P}^X \) is the distribution of a real-valued random variable \( X \), against general alternatives, \[7\] used that (1.1) can be untied from the class of test functions with the help of the so-called zero-bias transformation introduced in \[15\]. To be specific, a real-valued random variable \( X^* \) is said to have the \( X \)-zero-bias distribution if

\[ \mathbb{E}[f'(X^*)] = \mathbb{E}[Xf(X)] \]

holds for any of the respective test functions \( f \). If \( \mathbb{E}X = 0 \) and \( \mathbb{V}(X) = 1 \), the \( X \)-zero-bias distribution exists and is unique, and it has distribution function

\[ T^X(t) = \mathbb{E}[X(X - t) \mathbb{1}\{X \leq t\}], \quad t \in \mathbb{R}. \tag{1.3} \]

By (1.1), the standard Gaussian distribution is the unique fixed point of the transformation \( \mathbb{P}^X \mapsto \mathbb{P}^{X^*} \). Thus, the distribution of \( X \) is standard normal if, and only if,

\[ T^X = F. \tag{1.4} \]

In the spirit of characterization-based goodness-of-fit tests, an idea introduced in \[28\], this fixed point property directly admits a new class of testing procedures as follows. Letting \( \tilde{T}_n^X \) be an empirical version of \( T^X \) and \( \tilde{F}_n \) the empirical distribution function, both based on the standardized sample, \[7\] proposed a test for (1.2) based on the statistic

\[ G_n = n \int_{\mathbb{R}} \left| \tilde{T}_n^X(t) - \tilde{F}_n(t) \right|^2 w(t) \, dt, \]

where \( w \) is an appropriate weight function, which, in view of (1.4), rejects the normality hypothesis for large values of \( G_n \). As these tests have several desirable properties such as the consistency against general alternatives, and since they show a very promising performance in simulations, we devote this work to the question to what extent the fixed point property and the class of goodness-of-fit procedures may be generalized to other distributions.

Naturally, interest in applying Stein’s method to other distributions has already grown and delivered some corresponding results. Characterizations like (1.1) have been established en
mass (for an overview on characterizing Stein operators and further references, we recommend [26]). Charles Stein himself presented some ideas fundamental to the so-called density approach (see [36], chapter VI, and [37], section 5) which we shall use as the basis of our considerations. Related results for the special case of exponential families were already given in [23] and [34].

Another approach pioneered by [2] (see also [17]) includes working with the generator of the semi-group of operators corresponding to a Markov process whose stationary distribution is the one in consideration. A third advance is based on fixed point properties of probability transformations like the zero-bias transformation. Very general distributional transformations were introduced in [16] and refined in [13]. In the latter the transformations, and with them the explicit formulae, rely heavily on sign changes of the so-called biasing functions. These sign changes depend on the parameters of the distribution in consideration which renders the explicit representations impractical for the use in goodness-of-fit testing.

The starting point of the present paper is the density approach identity. Here, a result more general than (1.1) is provided by showing that, for suitable density functions $p$, a given real-valued random variable $X$ has density $p$ if, and only if,

$$E \left[ f'(X) + \frac{p'(X)}{p(X)} f(X) \right] = 0$$

holds for a sufficiently large class of test functions. We provide fixed point characterizations like (1.4) by using the analogy between (1.5) and (1.1) to extrapolate the explicit formula (1.3) of the zero-bias transformation to other distributions. Using this approach, these transformations will no longer be probability transformations, but we maintain the characterizing identity which suffices for the use in goodness-of-fit testing. Our confidence in the approach is manifested by the fact that it has already been implemented in [6] for the special case of the Gamma distribution.

The paper is organized as follows. We first introduce an appropriate setting for our considerations. In particular, we state the conditions for a density function to fit into our framework and prove the identity (1.5) in this specific setting. We then give our characterization results, distinguishing between distributions supported by the whole real line, those with semi-bounded support and distributions with bounded support. Throughout, we give examples of density functions of different nature to show that our conditions are not restrictive as well as to provide connections to characterizations that are already known and included in our statements. Then
we consider applications in goodness-of-fit testing and show that the proposed tests include
the classical Kolmogorov-Smirnov-, Anderson-Darling- and Cramér-von Mises procedures as
well as three modern tests considered in the literature.

2 Characterizations of distribution functions.

Throughout this section, let \((\Omega, A, \mathbb{P})\) be a probability space and \(p\) a non-negative den-
sity function supported by an interval \(\text{spt}(p) = [L, R]\), where \(-\infty \leq L < R \leq \infty\) and
\(\int_{L}^{R} p(x) \, dx = 1\). Denoting by \(P\) the distribution function associated with \(p\), we state the
following regularity conditions:

(C1) \(p\) is continuously differentiable on \((L, R)\),

(C2) \(p(x) > 0\) for every \(x \in (L, R)\),

(C3) for \(\kappa_p(x) = \left| \frac{p'(x) \min\{P(x), 1-P(x)\}}{p(x)} \right|\) we have \(\sup_{x \in (L, R)} \kappa_p(x) < \infty\),

(C4) \(\int_{L}^{R} (1 + |x|) |p'(x)| \, dx < \infty\),

(C5) \(\lim_{x \searrow L} \frac{P(x)}{p(x)} = 0\) and

(C6) \(\lim_{x \nearrow R} \frac{1-P(x)}{p(x)} = 0\).

Often, (C4) can be replaced by a weaker condition, and we include remarks on this at the
right places. First, we restate the Stein characterization for such density functions. We will
roughly follow the line of proof in \([27]\), but put more focus on the technical details that arise
from our special choice of test functions.

Definition 2.1. For a density function \(p\) with \(\text{spt}(p) = [L, R]\) that satisfies (C1) – (C3) we
denote by \(\mathcal{F}_p\) the set of all functions \(f : (L, R) \rightarrow \mathbb{R}\) that are differentiable on \((L, R)\), satisfy
\(\lim_{x \searrow L} f(x) p(x) = \lim_{x \nearrow R} f(x) p(x) = 0\),
and for which \(x \mapsto \frac{p'(x)}{p(x)} f(x)\) and \(x \mapsto f'(x)\) are uniformly bounded on the interval \((L, R)\).

In the following, we denote by \(\mathcal{L}^1\) the Lebesgue measure on \(\mathbb{R}\) and write \(X \sim p\mathcal{L}^1\) when a
random variable \(X\) has Lebesgue density \(p\).
Lemma 2.2. If $p$ is a density function with $\text{spt}(p) = [L, R]$ that satisfies (C1) – (C3) and $X : \Omega \to (L, R)$ is an arbitrary random variable, then $X \sim p\mathcal{L}^1$ if, and only if,

$$\mathbb{E} \left[ f'(X) + \frac{p'(X)}{p(X)} f(X) \right] = 0$$

for each $f \in \mathcal{F}_p$.

Proof. Note that if $X \sim p\mathcal{L}^1$, any $f \in \mathcal{F}_p$ satisfies

$$\mathbb{E} \left[ f'(X) + \frac{p'(X)}{p(X)} f(X) \right] = \int_{L}^{R} (f \cdot p)'(x) \, dx = \lim_{x \nearrow R} f(x) \cdot p(x) - \lim_{x \searrow L} f(x) \cdot p(x) = 0.$$

Here, we applied the fundamental theorem of calculus utilizing that the function

$$(L, R) \ni x \mapsto (f \cdot p)'(x) = f'(x) \cdot p(x) + \frac{p'(x)}{p(x)} f(x) \cdot p(x)$$

is integrable. For the converse, fix $t \in (L, R)$ and define $f^p_t : (L, R) \to \mathbb{R}$ through

$$f^p_t(x) = \frac{1}{p(x)} \int_{L}^{x} \left( \mathbb{1}_{(L,t]}(s) - P(t) \right) p(s) \, ds.$$

Apparently,

$$\lim_{x \nearrow R} f^p_t(x) \cdot p(x) = \int_{L}^{R} \left( \mathbb{1}_{(L,t]}(s) - P(t) \right) p(s) \, ds = P(t) - P(t) = 0,$$

and noting that $f^p_t(x) = \frac{1}{p(x)} P(x)(1 - P(t))$ for $x < t$, we also have $\lim_{x \searrow L} f^p_t(x) \cdot p(x) = 0$. Moreover, $f^p_t$ is differentiable with

$$f^{p'}_t(x) = -\frac{p'(x)}{p(x)} f^p_t(x) + \mathbb{1}_{(L,t]}(x) - P(t), \quad x \in (L, R). \quad (2.1)$$

With (C3) we get

$$\sup_{x \in (L,R)} \left| \frac{p'(x)}{p(x)} f^p_t(x) \right| \leq \sup_{x \in (L,R)} \left| \frac{p'(x) \min\{P(x), 1 - P(x)\}}{p^2(x)} \right| < \infty,$$

and (2.1) suffices to conclude that $f^p_t \in \mathcal{F}_p$. Thus, the assumption and (2.1) yield

$$0 = \mathbb{E} \left[ f^{p'}_t(X) + \frac{p'(X)}{p(X)} f^p_t(X) \right] = \mathbb{P}(X \leq t) - P(t).$$

As $t$ was arbitrary, $X \sim p\mathcal{L}^1$. \qed

Remark 2.3. Notice that the proof depends on $\mathcal{F}_p$ solely through $f^p_t$, and we may assume the functions in $\mathcal{F}_p$ to satisfy any property we are able to verify for $f^p_t$. For instance, condition
(C3) implies the boundedness of $\frac{p'}{p}f''$ and $f''$, and thus, $\frac{p'}{p}f$ and $f'$ are bounded for any function $f \in \mathcal{F}_p$. This fact trivially implies that the expectation figuring in the lemma exists. However, (C3) is not necessary to guarantee its existence. Since $f''(x) + \frac{p'(x)}{p(x)}f'(x)$, $x \in (L, R)$, is uniformly bounded by equation (2.1), we can directly assume that, for each $f \in \mathcal{F}_p$, the function $f' + \frac{p'}{p}f$ is integrable with respect to any probability measure. One might argue at this point that we could also require (possibly instead of (C3))

$$\sup_{x \in (L, R)} \left| \frac{\min\{P(x), 1 - P(x)\}}{p(x)} \right| < \infty.$$  

Indeed, this condition would allow us to claim that each function in $\mathcal{F}_p$ is bounded, but such a bound will not be enough for our considerations later on. Note that these assumptions are commonly made in the context of Stein’s method to get bounds on the solution of the Stein equation (see, e.g., Lemma 13.1 in [11]).

**Remark 2.4.** For later use, we note that if (C5) holds, any function $f \in \mathcal{F}_p$ is subject to

$$\lim_{x \searrow L} f(x) = 0$$

since $f''$ from the proof satisfies

$$\lim_{x \searrow L} f''(x) = \lim_{x \searrow L} \frac{1}{p(x)} \left( P(x)(1 - P(x)) \right) = 0.$$  

By analogy, if (C6) holds, then

$$\lim_{x \nearrow R} f''(x) = \lim_{x \nearrow R} \frac{1}{p(x)} \left( P(t)(1 - P(x)) \right) = 0,$$

and each function $f \in \mathcal{F}_p$ satisfies $\lim_{x \nearrow R} f(x) = 0$.

In a different form, the characterization given by Lemma 2.2 has successfully been applied for distributional approximations in the Curie-Weiss model (see [10]) or the hitting times of Markov chains (see [32]). For an overview, we refer to section 13 in [11]. In this paper, however, we use the characterization to derive another, more explicit identity that typifies distributions with density functions as above. We thereby generalize the fixed point properties of the well-known zero-bias and equilibrium transformations but also classical identities, such as the characterization of the exponential distribution through the mean residual life function.
2.1 Distributions supported by the real line.

In this subsection, we assume that $p : \mathbb{R} \rightarrow [0, \infty)$ is a probability density function supported by the whole real line, i.e., $\text{spt}(p) = \mathbb{R}$.

**Theorem 2.5.** Let $p$ be a density function with $\text{spt}(p) = \mathbb{R}$ that satisfies (C1) – (C4). Moreover, let $X : \Omega \rightarrow \mathbb{R}$ be an arbitrary random variable with distribution function $F$ and

$$
\mathbb{E} \left| \frac{p'(X)}{p(X)} \right| < \infty, \quad \mathbb{E} \left| \frac{p'(X)}{p(X)} X \right| < \infty. \tag{2.2}
$$

If $T^X_p : \mathbb{R} \rightarrow \mathbb{R}$ is defined by

$$
T^X_p(t) = \mathbb{E} \left[ -\frac{p'(X)}{p(X)} (X-t) \mathbb{1}\{X \leq t\} \right], \quad t \in \mathbb{R},
$$

then $X \sim p\mathcal{L}^1$ if, and only if, $T^X_p(t) = F(t)$ for every $t \in \mathbb{R}$.

**Proof.** Suppose first that $X \sim p\mathcal{L}^1$. Using (C4), we have

$$
\mathbb{E} \left[ \frac{p'(X)}{p(X)} \mathbb{1}\{X \leq t\} \right] = \int_{-\infty}^{t} p'(x) \, dx = p(t), \quad t \in \mathbb{R},
$$

and Tonelli’s theorem gives

$$
\int_{-\infty}^{t} \mathbb{E} \left[ \frac{p'(X)}{p(X)} \mathbb{1}\{X \leq s\} \right] \, ds = \int_{-\infty}^{t} \int_{-\infty}^{s} |p'(x)| \, dx \, ds = \int_{-\infty}^{t} |p'(x)| (t-x) \, dx < \infty
$$

for every $t \in \mathbb{R}$. Now, Fubini’s theorem yields

$$
F(t) = \int_{-\infty}^{t} p(s) \, ds = \int_{-\infty}^{t} \mathbb{E} \left[ \frac{p'(X)}{p(X)} \mathbb{1}\{X \leq s\} \right] \, ds
$$

$$
= \mathbb{E} \left[ \frac{p'(X)}{p(X)} \int_{-\infty}^{t} \mathbb{1}\{X \leq s\} \, ds \right]
$$

$$
= T^X_p(t), \quad t \in \mathbb{R}.
$$

For the converse part, we assume $T^X_p(t) = F(t), t \in \mathbb{R}$. Putting

$$
d^X_p(t) = \mathbb{E} \left[ \frac{p'(X)}{p(X)} \mathbb{1}\{X \leq t\} \right], \quad t \in \mathbb{R},
$$

condition (2.2) entails

$$
\mathbb{E} \left[ \int_{-\infty}^{t} \frac{|p'(X)|}{p(X)} \mathbb{1}\{X \leq s\} \, ds \right] = \mathbb{E} \left[ \frac{|p'(X)|}{p(X)} (t-X) \mathbb{1}\{X \leq t\} \right] < \infty
$$
for every $t \in \mathbb{R}$. Thus, Fubini’s theorem implies
\[
\int_{-\infty}^{t} d_{p}^{X}(s) \, ds = T^{X}_{p}(t), \quad t \in \mathbb{R}.
\]
From this relation we infer
\[
\int_{\mathbb{R}} d_{p}^{X}(s) \, ds = \lim_{t \to \infty} T^{X}_{p}(t) = \lim_{t \to \infty} F(t) = 1,
\]
and since $T^{X}_{p}$ is increasing, it follows that $d_{p}^{X} \geq 0$ $\mathcal{L}^{1}$-almost everywhere. Hence, $d_{p}^{X}$ is the density function of $T^{X}_{p} = F$. Using the first part of (2.2), dominated convergence gives
\[
\mathbb{E} \left[ p'(X) \right] = \lim_{t \to \infty} \mathbb{E} \left[ \frac{p'(X)}{p(X)} 1_{\{X \leq t\}} \right] = \lim_{t \to \infty} d_{p}^{X}(t) = 0.
\]
For each $f \in \mathcal{F}_{p}$ we conclude
\[
\mathbb{E} \left[ f'(X) \right] = \int_{\mathbb{R}} f'(s) d_{p}^{X}(s) \, ds
\]
\[
= \int_{-\infty}^{0} f'(s) \mathbb{E} \left[ \frac{p'(X)}{p(X)} 1_{\{X \leq s\}} \right] \, ds + \int_{0}^{\infty} f'(s) \mathbb{E} \left[ \frac{p'(X)}{p(X)} 1_{\{X > s\}} \right] \, ds
\]
\[
= \mathbb{E} \left[ \frac{p'(X)}{p(X)} (f(0) - f(X)) 1_{\{X \leq 0\}} \right] + \mathbb{E} \left[ \frac{p'(X)}{p(X)} (f(X) - f(0)) 1_{\{X > 0\}} \right]
\]
\[
= \mathbb{E} \left[ \frac{p'(X)}{p(X)} f(X) \right].
\]
Here, Fubini’s theorem is applicable since $f'$ is uniformly bounded and we have the second part of (2.2). Lemma 2.2 yields the claim. \hfill \Box

Remark 2.6. Theorem 2.5 characterizes a distribution as the unique fixed point of the map $F \mapsto T^{X}_{p}$ only within the class of all distributions that satisfy (2.2). Assumption (C4) is equivalent to claiming that (2.2) is satisfied for the density function $p$ itself (i.e., if $X \sim p\mathcal{L}^{1}$).

We consider two examples to see what the validation of our conditions looks like. In each case, (C1) and (C2) are satisfied trivially.

Example 2.7 (Standard Cauchy Distribution). Let $p(x) = \pi^{-1}(1 + x^{2})^{-1}$, $x \in \mathbb{R}$. We easily see that
\[
\int_{\mathbb{R}} (1 + |x|)|p'(x)| \, dx = 2 \int_{\mathbb{R}} \frac{|x| + x^{2}}{\pi(1 + x^{2})^{2}} \, dx
\]
\[
= -\frac{1}{\pi} \int_{-\infty}^{0} \frac{2x}{(1 + x^{2})^{2}} \, dx + \frac{1}{\pi} \int_{0}^{\infty} \frac{2x}{(1 + x^{2})^{2}} \, dx + 2 \int_{\mathbb{R}} \frac{x^{2} p(x)}{1 + x^{2}} \, dx
\]
\[
\leq 2(\pi^{-1} + 1),
\]
which is (C4). By L’Hospital’s rule, we have
\[
\lim_{x \to -\infty} \kappa_p(x) \leq \lim_{x \to -\infty} 2|x| \left( \int_{-\infty}^{x} \frac{1}{1+s^2} \, ds \right) = \lim_{x \to -\infty} \frac{2x^2}{1+x^2} = 2
\]
and, by symmetry, \(\lim_{x \to \infty} \kappa_p(x) \leq 2\). Together with the continuity of \(\kappa_p\), (C3) follows. Thus, Theorem 2.5 holds for the Cauchy distribution. In this case,
\[
\frac{p'(x)}{p(x)} = -\frac{2x}{1+x^2}, \quad x \in \mathbb{R},
\]
and (2.2) reads as \(\mathbb{E}[(1 + X^2)^{-1}|X|] < \infty\).

**Example 2.8** (Mean-zero Gaussian Distribution). For \(x \in \mathbb{R}\) let
\[
p(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{x^2}{2\sigma^2}\right),
\]
where \(0 < \sigma^2 < \infty\). Condition (C4) follows from the existence of mean and variance of the normal distribution. Using the readily verified identities
\[
1 - P(x) \leq \frac{\sigma^2}{x} p(x), \quad x > 0, \quad \text{and} \quad P(x) \leq -\frac{\sigma^2}{x} p(x), \quad x < 0,
\]
we see that \(\kappa_p(x) \leq 1\) for every \(x \in \mathbb{R}\). Therefore, (C3) and, hence, Theorem 2.5 hold. Here, \(\frac{p'(x)}{p(x)} = -\frac{x}{\sigma^2}\) and (2.2) is satisfied if \(\mathbb{E}X^2 < \infty\). The function \(T^X_p\) takes the form
\[
T^X_p(t) = \sigma^{-2} \mathbb{E}\left[X(X-t) \mathbb{1}\{X \leq t\}\right], \quad t \in \mathbb{R}. \tag{2.3}
\]
This is the explicit formula of the distribution function of the zero-bias transformation of \(\mathbb{P}X\), so the example at hand needs further examination. To that end, we recall that the authors of [15] showed that for any random variable \(X\) with mean zero and variance \(\sigma^2\) there is a unique distribution for the random variable \(X^*\) such that
\[
\sigma^2 \mathbb{E}[f'(X^*)] = \mathbb{E}[Xf(X)]
\]
for each absolutely continuous function \(f\) for which the expectation on the right-hand side exists. They provided the explicit formula (2.3) for the distribution function of this so-called zero-bias distribution. The transformation \(\mathbb{P}X \to \mathbb{P}X^*\) has the normal distribution \(\mathcal{N}(0, \sigma^2)\) as its unique fixed point and typifies this distribution within the class of all distributions with mean zero and variance \(\sigma^2\). Through Theorem 2.5, we generalize the characterization result to the class of all distributions with \(\mathbb{E}X^2 < \infty\), but if \(\mathbb{E}X \neq 0\) or \(\forall(X) \neq \sigma^2\), \(T^X_p\) may no longer be a distribution function.
Remark 2.9. Putting further restrictions on the distribution of $X$ in Theorem 2.5 such that $d^X_p$ is a probability density function, we have actually shown the existence of a distribution for some random variable $X_p$ with

\[ E[f'(X_p)] = E\left[-\frac{p'(X)}{p(X)} f(X)\right] \]

for each $f \in \mathcal{F}_p$, and we think of $\mathbb{P}^X \mapsto T^X_p$ as a distributional transformation. These additional restrictions (for the normal distribution they are $E X = 0$ and $\var{X} = \sigma^2$) are very strong and scale down the class of distributions in which the characterization holds. Therefore, we do not cling on to distributional transformations but extract whichever information we can get from the explicit formula of $T^X_p$.

### 2.2 Distributions with semi-bounded support.

First, we show that a result similar to Theorem 2.5 holds for density functions whose support is bounded from below. Namely, we let $p : \mathbb{R} \to [0, \infty)$ be a probability density function with $\text{spt}(p) = [L, \infty)$, $L > -\infty$. The most important case is $L = 0$, that is, density functions supported by the positive half line.

**Theorem 2.10.** Assume that $p$ is a density function with $\text{spt}(p) = [L, \infty)$ that satisfies the conditions (C1) – (C5). Let $X : \Omega \to (L, \infty)$ be a random variable with distribution function $F$ and

\[ E \left| \frac{p'(X)}{p(X)} \right| < \infty \quad \text{(only if $L \neq 0$),} \quad E \left\| \frac{p'(X)}{p(X)} X \right\| < \infty. \quad (2.4) \]

Define $T^X_p : \mathbb{R} \to \mathbb{R}$ by

\[ T^X_p(t) = E \left[ -\frac{p'(X)}{p(X)} \left( \min\{X, t\} - L \right) \right], \quad t \in (L, \infty), \]

and $T^X_p(t) = 0$, $t \in (-\infty, L]$. Then $X \sim p\mathcal{L}^1$ if, and only if, $T^X_p(t) = F(t)$ for every $t \in \mathbb{R}$.

**Proof.** If $X \sim p\mathcal{L}^1$, we have

\[ E \left[ -\frac{p'(X)}{p(X)} \mathbb{1}_{\{X > t\}} \right] = - \int_t^\infty p'(x) \, dx = p(t), \quad t \in (L, \infty), \]

10
and

\[ F(t) = \int L \, p(s) \, ds = \int L \, E \left[ \frac{-p'(X)}{p(X)} \mathbb{1}\{X > s\} \right] \, ds \]

\[ = E \left[ \frac{-p'(X)}{p(X)} \int L \, \mathbb{1}\{X > s\} \, ds \right] \]

\[ = T^X_p(t), \quad t \in (L, \infty). \]

The arguments on the integrability that render these calculations valid are similar to those in the proof of Theorem 2.5. For the converse part, we put

\[ d^X_p(t) = E \left[ \frac{-p'(X)}{p(X)} \mathbb{1}\{X > t\} \mathbb{1}_{(L, \infty)}(t), \quad t \in \mathbb{R}, \right] \]

and notice that (2.4) implies

\[ E \left[ \int L \, \frac{|p'(X)|}{p(X)} \mathbb{1}\{X > s\} \, ds \right] = E \left[ \frac{|p'(X)|}{p(X)} \left( \min\{X, t\} - L \right) \right] \]

\[ \leq E \left[ \frac{p'(X)}{p(X)} X \right] + |L| \cdot E \left[ \frac{p'(X)}{p(X)} \right] \]

\[ < \infty \]

for every \( t \in (L, \infty) \). Thus, Fubini’s theorem gives

\[ \int L \, d^X_p(s) \, ds = T^X_p(t), \quad t \in \mathbb{R}. \]

We conclude that \( d^X_p \) is the density function of \( T^X_p = F \), and with \( \lim_{x \to L} f(x) = 0 \) (cf. Remark 2.4) we get

\[ E[f'(X)] = \int \mathbb{R} f'(s) d^X_p(s) \, ds = \int L \, \mathbb{E} \left[ \frac{-p'(X)}{p(X)} \mathbb{1}\{X > s\} \right] \, ds \]

\[ = E \left[ \frac{p'(X)}{p(X)} \int L \, f'(s) \, ds \right] \]

\[ = E \left[ \frac{p'(X)}{p(X)} f(X) \right] \]

for each \( f \in \mathcal{F}_p \). By Lemma 2.2 we are done. \( \square \)

**Example 2.11 (Standard Lévy Distribution).** For \( x > 0 \) let

\[ p(x) = \frac{1}{\sqrt{2\pi}} x^{-3/2} \exp \left( -\frac{1}{2x} \right). \]
As explained in Remark 2.12, the fact that
\[ \int_0^\infty |p'(x)| \, dx = \int_0^{1/3} p'(x) \, dx - \int_{1/3}^\infty p'(x) \, dx = 2 \cdot p(1/3) < \infty \]
is enough to cover condition (C4). By a change of variable,
\[ P_X(x) = 2x^{3/2} \exp \left( \frac{1}{2x} \right) \int_{1/\sqrt{x}}^\infty \exp \left( -\frac{t^2}{2} \right) \, dt \leq 2x^2. \]

Therefore, (C5) holds and we conclude that \( \lim_{x \to 0} \kappa_p(x) \leq 1 \). Finally, the estimate
\[ \frac{1}{1 - P_X(x)} = 2x^{3/2} \exp \left( \frac{1}{2x} \right) \int_{1/\sqrt{x}}^\infty \exp \left( -\frac{t^2}{2} \right) \, dt \leq 2 \exp \left( \frac{1}{2x} \right) \]
yields \( \lim_{x \to \infty} \kappa_p(x) \leq 3 \), and (C3) is valid. We have
\[ \frac{p'(x)}{p(x)} = \frac{1}{2x^2} - \frac{3}{2x}, \quad x > 0, \]
and Theorem 2.10 holds, with (2.4) reading as \( \mathbb{E}X^{-1} < \infty \).

In the following, we discuss the conditions we impose on the density functions in consideration.

Remark 2.12. We cannot attenuate the conditions (C1), (C2) and (C5) in a meaningful way. Yet, to justify the use of Fubini’s theorem in the necessity part of Theorem 2.10 we argue (in contrast to Theorem 2.5) that
\[ \int_L^t \mathbb{E} \left[ \frac{|p'(X)|}{p(X)} \mathbbm{1}\{X > s\} \right] \, ds = \int_L^t \int_s^\infty |p'(x)| \, dx \, ds \leq (t - L) \int_L^\infty |p'(x)| \, dx \]
for every \( t \in (L, \infty) \), and only require \( \int_L^\infty |p'(x)| \, dx < \infty \) instead of (C4) (used in Example 2.11). Alternatively, in the case \( L = 0 \), we have
\[ \int_0^\infty \mathbb{E} \left[ \frac{|p'(X)|}{p(X)} \mathbbm{1}\{X > s\} \right] \, ds = \int_0^\infty \int_s^\infty |p'(x)| \, dx \, ds \leq \int_0^\infty x |p'(x)| \, dx, \]
so (C4) may be replaced with \( \int_0^\infty x |p'(x)| \, dx < \infty \) and \( \int_t^\infty |p'(x)| \, dx < \infty, \quad t > 0 \). This last case is particularly useful for density functions with \( \lim_{x \to 0} p(x) = \infty \), for which the integral \( \int_0^\infty |p'(x)| \, dx \) might not exist (cf. Example 2.13 below).

Example 2.13 (Weibull distribution). For \( \lambda, k > 0 \) let
\[ p(x) = \lambda^{-k} k x^{k-1} \exp \left( -\lambda^{-k} x^k \right), \quad x > 0. \]
From the inequalities

\[
\frac{1 - P(x)}{p(x)} \leq \frac{\lambda^k}{x^{k-1}} \quad \text{and} \quad \frac{P(x)}{p(x)} \leq \frac{x \cdot \exp(\lambda^{-k}x^k)}{k},
\]

we infer that (C3) and (C5) are valid. By Remark 2.12 to cover (C4) it suffices to know that

\[
\int_t^\infty |p'(x)| \, dx \leq \frac{|k-1|}{t} + \frac{k}{\lambda^k} \int_0^\infty x^{k-1} p(x) \, dx < \infty
\]

for every \( t \in (0, \infty) \), and

\[
\int_0^\infty x |p'(x)| \, dx \leq |k-1| + \frac{k}{\lambda^k} \int_0^\infty x^k p(x) \, dx < \infty.
\]

**Remark 2.14.** The purpose of condition (C3) is to assure the boundedness of \( f' \), \( f \in \mathcal{F}_p \), which, in turn, allows us to use Fubini’s theorem in the last step of the proof of our characterization. It is conceivable to replace (C3) with the requirement that

\[
x \mapsto \frac{p'(x)}{p^2(x)} \min \{ P(x), 1 - P(x) \}
\]

is integrable on \((L, \infty)\) which would entail the integrability of \( f' \), \( f \in \mathcal{F}_p \). In this case, Theorem 2.10 holds if we only require \( E \left[ \frac{|p'(X)|}{p(X)} \right] < \infty \) and drop the second integrability condition. However, even for simple density functions (e.g., the exponential distribution density) the function in (2.5) is not integrable.

We give another example with connections to characterizations and transformations already discussed in the literature.

**Example 2.15** (Exponential Distribution). Let \( p(x) = \lambda \exp(-\lambda x) \), \( x > 0 \), with rate parameter \( \lambda > 0 \). Here, (C3) is satisfied since \( \kappa_p(x) \leq 1 \) for every \( x > 0 \). Condition (C4) holds since

\[
\int_0^\infty (1 + x) |p'(x)| \, dx = \lambda + \lambda^2,
\]

and (C5) is verified as

\[
\lim_{x \downarrow 0} \frac{P(x)}{x \downarrow 0} = \lim_{x \downarrow 0} \frac{1}{\lambda} ( \exp(\lambda x) - 1 ) = 0.
\]

We have \( \frac{p'(x)}{p(x)} = -\lambda \), and the integrability condition \( (2.5) \) becomes \( E X < \infty \). In this case, Theorem 2.10 reads as follows. A positive random variable \( X \) with distribution function \( F \) and \( E X < \infty \) follows an exponential law if, and only if,

\[
E[X] F(t) = E \left[ \min \{ X, t \} \right], \quad t > 0.
\]
As [1] noted, this identity is equivalent to the well-known characterization of the exponential distribution via the mean residual life function, which states that a positive random variable \( X \) with \( \mathbb{E}X < \infty \) follows an exponential law if, and only if,

\[
\mathbb{E}[X - t \mid X > t] = \mathbb{E}[X], \quad t > 0.
\]

For yet another observation, assume that \( X \) is a positive random variable with \( \mathbb{E}X = \lambda - 1 \). With \( d_p^X \) as in the proof of Theorem 2.10, we have \( d_p^X \geq 0 \) and

\[
\int_0^\infty d_p^X(t) \, dt = \lambda \int_0^\infty \mathbb{P}(X > t) \, dt = \lambda \mathbb{E}X = 1.
\]

If \( X^e \) is a random variable with density function \( d_p^X \), the proof of Theorem 2.10 shows that

\[
\mathbb{E}[f'(X^e)] = \lambda \mathbb{E}[f(X)]
\]

for each \( f \in F_p \). Up to a change in the class of test functions, this last line is the defining equation of the equilibrium distribution with respect to \( X \). Lemma 2.2 implies that when restricting to \( \mathbb{E}X = \lambda^{-1} \), the exponential distribution with parameter \( \lambda \) is the unique fixed point of the equilibrium transformation \( \mathbb{P}^X \mapsto \mathbb{P}^{X^e} \). This fact is used for approximation arguments with Stein’s method (see, for instance, [32], who introduced the equilibrium distribution, as well as chapter 13.4 in [11] and section 5 in [35]). As in the case of the zero-bias transformation, we have generalized this characterization in the sense that the explicit formula of the equilibrium distribution, given by \( T_p^X \), uniquely identifies the exponential distribution with parameter \( \lambda \) within the class of all distributions \( \mathbb{P}^X \) with \( \mathbb{E}X < \infty \). Note that for random variables with \( \mathbb{E}X \neq \lambda^{-1} \), \( d_p^X \) is not a probability density function and, hence, \( T_p^X \) is not a distribution function.

**Remark 2.16.** If \( p : \mathbb{R} \to [0, \infty) \) is a probability density function supported by the whole real line that satisfies (C1), (C2) and (C4), then

\[
\mathbb{E} \left[ \frac{p'(X)}{p(X)} \right] = \int_{\mathbb{R}} p'(x) \, dx = 0 \quad (2.6)
\]

if \( X \sim p\mathcal{L}^1 \). In fact, we have seen in the proof of Theorem 2.5 that if \( X \) satisfies (2.2), the assumption \( T_p^X = F \) also leads to (2.6). Thus, we have

\[
\mathbb{E} \left[ \frac{p'(X)}{p(X)} \mathbbm{1}\{X \leq t\} \right] = \mathbb{E} \left[ -\frac{p'(X)}{p(X)} \mathbbm{1}\{X > t\} \right], \quad t \in \mathbb{R}.
\]
In Theorem 2.10 this needs not to be true. Since we only used the right-hand side of the above equation for the proof of Theorem 2.10, we could try to go through that proof taking the left-hand side as $d_{p}^{X}$. Indeed, requiring that the limit $\lim_{x \to L} p(x)$ exists, slight adaptations (as we will make them in section 2.3) preserve all arguments except for the very last calculation, where it fails on the unboundedness of the support.

To complete this subsection, we state the characterization result for a probability density function $p$ with $\text{spt}(p) = (-\infty, R]$, $R < \infty$. The proof runs along the lines of the proofs of Theorem 2.5 and Theorem 2.10 combining methods from both.

**Corollary 2.17.** Let $p$ be a density function with $\text{spt}(p) = (-\infty, R]$ that satisfies (C1) – (C4) and (C6). Moreover, let $X : \Omega \to (-\infty, R)$ be a random variable with distribution function $F$ and

$$
E \left[ \frac{p'(X)}{p(X)} \right] < \infty \ (\text{only if } R \neq 0), \quad E \left| \frac{p'(X)}{p(X)} X \right| < \infty.
$$

Then $X \sim p \mathcal{L}^{1}$ if, and only if, $T_{p}^{X}(t) = F(t)$ for every $t \in \mathbb{R}$, where

$$
T_{p}^{X}(t) = E \left[ -\frac{p'(X)}{p(X)} (X - t) 1\{X \leq t\} \right], \quad t \in (-\infty, R),
$$

and $T_{p}^{X}(t) = 1$, $t \in [R, \infty)$.

**Remark 2.18.** Again, we do not need condition (C4) to full extent. Even though we cannot handle it as flexible as in the case of Theorem 2.10 (cf. Remark 2.12), the existence of the integrals

$$
\int_{-\infty}^{t} (1 + |x|) |p'(x)| \, dx, \quad t \in (-\infty, R),
$$

covers all required implications of (C4).

### 2.3 Distributions with bounded support.

It remains to study probability density functions $p : \mathbb{R} \to [0, \infty)$ with $\text{spt}(p) = [L, R]$, where $L > -\infty$ and $R < \infty$. The first observation in the proofs of Theorem 2.5 and Theorem 2.10 relies on the fact that the limit to $R(= \infty)$ of the density function is zero. For densities with bounded support we would have to state this property as a restrictive prerequisite. Therefore, we adapt our arguments in order to weaken this specific presumption. Namely, we assume that the limit of the density function to one endpoint of the support exists.
Theorem 2.19. Assume that \( p \) is a density function with \( \text{spt}(p) = [L, R] \) that satisfies the conditions (C1) – (C6), and for which the limit \( \lim_{x \searrow R} p(x) \) exists. Let \( X : \Omega \to (L, R) \) be a random variable with distribution function \( F \) and

\[
E \left[ \frac{|p'(X)|}{p(X)} \right] < \infty. \tag{2.7}
\]

Moreover, let

\[
T_p^X(t) = E \left[ -\frac{p'(X)}{p(X)} \left( \min\{X, t\} - L \right) \right] + (t - L) \lim_{x \searrow R} p(x), \quad t \in (L, R),
\]

\( T_p^X(t) = 0, \quad t \in (-\infty, L] \), and \( T_p^X(t) = 1, \quad t \in [R, \infty) \). Then \( X \sim p \mathcal{L}^1 \) if, and only if, \( T_p^X(t) = F(t) \) for every \( t \in \mathbb{R} \).

Proof. Assuming \( X \sim p \mathcal{L}^1 \), (C4) allows for the calculation

\[
E \left[ -\frac{p'(X)}{p(X)} \mathbb{1}\{X > t\} \right] = -\int_t^R p'(x) \, dx = p(t) - \lim_{x \searrow R} p(x), \quad t \in (L, R).
\]

With Fubini’s theorem we get

\[
F(t) = E \left[ -\frac{p'(X)}{p(X)} \int_t^L \mathbb{1}\{X > s\} \, ds \right] + (t - L) \lim_{x \searrow R} p(x) = T_p^X(t)
\]

for every \( t \in (L, R) \). For the converse part, we put

\[
d_p^X(t) = \left( E \left[ -\frac{p'(X)}{p(X)} \mathbb{1}\{X > t\} \right] + \lim_{x \searrow R} p(x) \right) \mathbb{1}(L, R)(t), \quad t \in \mathbb{R}.
\]

Since we have

\[
E \left[ \int_t^L \frac{|p'(X)|}{p(X)} \mathbb{1}\{X > s\} \, ds \right] \leq (R - L) E \left[ \frac{|p'(X)|}{p(X)} \right] < \infty \tag{2.8}
\]

for every \( t \in (L, R) \), Fubini’s theorem yields that \( d_p^X \) is the density function of \( T_p^X = F \).

Finally, we conclude that

\[
E[f'(X)] = \int_L^R f'(s) \mathbb{1}\{X > s\} \, ds + \lim_{x \searrow R} p(x) \int_L^R f'(s) \, ds \]

\[
= E \left[ -\frac{p'(X)}{p(X)} \int_L^X f'(s) \, ds \right] + \lim_{x \searrow R} p(x) \left( \lim_{x \searrow R} f(x) - \lim_{x \searrow L} f(x) \right)
\]

\[
= E \left[ -\frac{p'(X)}{p(X)} f(X) \right]
\]

for each \( f \in \mathcal{F}_p \). \( \square \)
Remark 2.20. When \( \lim_{x \to R} p(x) = 0 \), Theorem 2.10 carries over without essential changes. Theorem 2.19 becomes relevant whenever \( \lim_{x \to R} p(x) \in (0, \infty) \), in which case we need both conditions, (C5) and (C6).

The reasoning from Remark 2.12 can be applied, i.e., we do not need condition (C4) to full extend. Again, \( \int_L^R |p'(x)| \, dx < \infty \) suffices, and in the case \( L = 0 \), the conditions \( \int_t^R |p'(x)| \, dx < \infty \), \( t \in (0, R) \), and \( \int_0^R x|p'(x)| \, dx < \infty \) can substitute (C4).

The boundedness of the support of \( p \) gives us some freedom concerning the integrability condition (2.7). Replacing (2.8) with

\[
E \left[ \int_L^t \frac{|p'(X)|}{p(X)} 1\{X > s\} \, ds \right] = E \left[ \frac{|p'(X)|}{p(X)} (X - L) \right],
\]

the weaker condition \( E \left[ \frac{p'(X)}{p(X)} X \right] < \infty \) compensates for (2.7) whenever \( L = 0 \).

Example 2.21 (Beta distribution). Let \( \alpha > 0, \beta > 1 \) and

\[
p(x) = B(\alpha, \beta)^{-1} x^{\alpha-1} (1 - x)^{\beta-1}, \quad x \in (0, 1),
\]

where \( B(\cdot, \cdot) \) denotes the Beta function. Using \( \beta > 1 \), we get

\[
\int_0^t x|p'(x)| \, dx \leq |\alpha - 1| + \alpha \quad \text{and} \quad \int_t^1 |p'(x)| \, dx \leq \frac{|\alpha - 1|}{\ell} + \alpha + \beta - 1
\]

for every \( t \in (0, 1) \), which ensures (C4). The estimates

\[
\frac{P(x)}{p(x)} \leq \alpha^{-1} x(1 - x)^{1-\beta} \quad \text{and} \quad \frac{1 - P(x)}{p(x)} \leq \beta^{-1} (1 - x) \max\{x^{1-\alpha}, 1\}
\]

imply that (C3) and (C5) hold. Technically, as \( \lim_{x \to 1} p(x) = 0 \), these distributions are included in Theorem 2.10. Nevertheless, we show them in the context of distributions with bounded support to make the following comments. Below, we state a similar result for cases where the limit to the left endpoint of the support of the density function exists but the limit to the right endpoint might not. The corresponding Corollary includes Beta distributions with \( \alpha > 1 \) and \( \beta > 0 \). This implies that whenever \( \alpha, \beta > 1 \), both characterization results hold (note that the problem discussed in Remark 2.16 does not arise). However, the Beta distribution also marks the limitation of our characterizations. If \( 0 < \alpha, \beta < 1 \), our results fail to hold since none of the required limits exist.

Example 2.22 (Uniform distribution). Let \( p(x) = (R - L)^{-1}, x \in (L, R) \), be the density function of the uniform distribution on the interval \((L, R)\). Since \( p'(x) = 0 \) for \( x \in (L, R) \),
(C3) and (C4) are trivially satisfied. The conditions (C5) and (C6) follow from

\[
\frac{P(x)}{p(x)} = x - L \quad \text{and} \quad \frac{1 - P(x)}{p(x)} = R - x.
\]

Theorem 2.19 reads as follows. A random variable \( X : \Omega \to (L, R) \) is distributed uniformly over \((L, R)\) if, and only if, its distribution function has the form

\[
F(t) = \begin{cases} 
0 & \text{for } t \leq L, \\
\frac{t-L}{R-L} & \text{for } t \in (L, R), \\
1 & \text{for } t \geq R.
\end{cases}
\]

Apparently, we recovered the observation that the explicitly calculable form of the uniform distribution function uniquely identifies this distribution.

We conclude this section by stating the last characterization identity, complementary to Theorem 2.19.

**Corollary 2.23.** Let \( p \) be a probability density function with \( \text{spt}(p) = [L, R] \) that satisfies the conditions (C1) – (C6), and for which \( \lim_{x \searrow L} p(x) \) exists. Further, let \( X : \Omega \to (L, R) \) be a random variable with distribution function \( F \) and

\[
\mathbb{E} \left[ \frac{|p'(X)|}{p(X)} \right] < \infty.
\]

Define the function \( T^X_p \) through

\[
T^X_p(t) = \mathbb{E} \left[ -\frac{p'(X)}{p(X)} (X - t) \mathbb{1}_{\{X \leq t\}} \right] + (t - L) \lim_{x \searrow L} p(x), \quad t \in (L, R),
\]

\( T^X_p(t) = 0, \; t \in (-\infty, L] \), and \( T^X_p(t) = 1, \; t \in [R, \infty) \). Then \( X \sim p\mathcal{L}^1 \) if, and only if, \( T^X_p(t) = F(t) \) for every \( t \in \mathbb{R} \).

The proof combines arguments from the proofs of Theorem 2.16 and Theorem 2.19. Similar to Remark 2.18, we can replace (C4) with

\[
\int_L^t |p'(x)||t - x| \, dx < \infty, \quad t \in (L, R),
\]

but even

\[
\int_L^R |p'(x)| \, dx < \infty
\]
suffices. By analogy to Remark 2.20 if \( R = 0 \), we can replace the integrability condition on \( X \) with
\[
E \left| \frac{p'(X)}{p(X)} X \right| < \infty.
\]

In general, the two characterization results for density functions with bounded support give a good handle on a wide variety of wrapped and truncated distributions.

### 2.4 A further result.

Let \( P \) be a distribution function whose density \( p \) satisfies the prerequisites for one of our characterization statements (Theorems 2.5, 2.10, 2.19 and Corollaries 2.17, 2.23). Let \( X \) be a random variable satisfying the proper integrability condition, with values in the interior of \( \text{spt}(p) \), and with distribution function \( F \). Up to now, we have shown that \( X \sim p \mathcal{L}^1 \), i.e., \( F = P \), holds if, and only if, \( T_p^X = F \). Next, we provide a sufficient condition such that \( T_p^X = P \) also implies \( X \sim p \mathcal{L}^1 \).

**Corollary 2.24.** Let \( p \) be a probability density function with \( \text{spt}(p) = [L, R] \), where \(-\infty \leq L < R \leq \infty \), that satisfies (C1) – (C4). Assume that (C5), (C6) hold and that \( \lim_{x \to R} p(x) \) or \( \lim_{x \to L} p(x) \) exist, if any of those conditions is required for the corresponding characterization result. Denote by \( P \) the distribution function associated with \( p \) and let \( X \) be a random variable with values in \((L, R)\) and
\[
E \left| \frac{p'(X)}{p(X)} X \right| < \infty \quad \text{and/or} \quad E \left[ \frac{|p'(X)|}{p(X)} \right] < \infty.
\]

Finally, assume that the distribution of \( Z \sim p \mathcal{L}^1 \) is uniquely determined by the values
\[
\left\{ E \left[ -\frac{p'(Z)}{p(Z)} f(Z) \right] \right\}_{f \in \mathcal{F}_p}.
\]

Then \( X \sim p \mathcal{L}^1 \) if, and only if, \( T_p^X = P \). (Depending on \( \text{spt}(p) \), \( T_p^X \) is defined as in the respective characterization result from the previous sections.)

**Proof.** If \( X \sim p \mathcal{L}^1 \), our preceding results yield \( T_p^X = F = P \). Conversely, if \( T_p^X = P \), it follows that \( d_p^X \) (defined in accordance with \( T_p^X \)) is the density function of \( T_p^X = P \), i.e., \( d_p^X = p \). Thus, any \( f \in \mathcal{F}_p \) satisfies
\[
E[f'(Z)] = \int_L^R f'(s) d_p^X(s) \, ds = E \left[ -\frac{p'(X)}{p(X)} f(X) \right].
\]
Lemma 2.2 gives
\[ E \left[ - \frac{p'(Z)}{p(Z)} f(Z) \right] = E \left[ - \frac{p'(X)}{p(X)} f(X) \right] \]
for each \( f \in \mathcal{F}_p \). By assumption, this implies \( X \sim pL^1 \).

\[ \blacklozenge \]

**Example 2.25** (Standard normal distribution). Let \( p(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \ x \in \mathbb{R} \), and \( Z \sim pL^1 \). We easily verify (e.g. by explicit constructions) that any monomial can be approximated with functions in \( \mathcal{F}_p \). Using dominated convergence, we derive from the values
\[ \left\{ E \left[ - \frac{p'(Z)}{p(Z)} f(Z) \right] \right\}_{f \in \mathcal{F}_p} = \left\{ E [Zf(Z)] \right\}_{f \in \mathcal{F}_p} \]
the sequence \( \{EZ^k\}_{k \in \mathbb{N}} \). Since the normal distribution is uniquely determined by its sequence of moments (see, for instance, Theorem 30.1 and Example 30.1 in \([8]\)), the condition in Corollary 2.24 is satisfied.

3 Applications to goodness-of-fit testing.

The idea to use distributional characterizations as a basis for statistics in goodness-of-fit testing problems is classic, see \([29]\) and \([31]\). In this spirit and regarding the results of the previous section, we propose goodness-of-fit tests for any distribution with a density function considered in section 2. To be specific, we consider a parametric family of distributions \( \mathcal{P} = \{p_\vartheta L^1 \mid \vartheta \in \Theta\} \), where we assume that \( p_\vartheta \) satisfies the conditions of one of our characterization identities. Here, \( \vartheta \in \Theta \), where \( \Theta \subset \mathbb{R}^d \) is open, \( d \in \mathbb{N} \), and \( \text{spt}(p_\vartheta) = [L, R] \) for fixed \( -\infty \leq L < R \leq \infty \).

Testing the fit of a random variable \( X \) to \( \mathcal{P} \) means to test the hypothesis
\[ H_0 : \mathbb{P}^X \in \mathcal{P} \quad (3.1) \]
against general alternatives. For the construction of suitable statistics, assume that the integrability conditions
\[ E \left| \frac{p'_\vartheta(X)}{p_\vartheta(X)} X \right| < \infty \quad \text{and/or} \quad E \left[ \left| \frac{p'_\vartheta(X)}{p_\vartheta(X)} \right| \right] < \infty \]
are satisfied, corresponding to the respective characterization result from the previous section.

Additionally, let \( s : (L, R) \times \Theta \to (L, R) \) be a measurable function, such that \( X \sim p_\vartheta L^1 \) if, and only if, \( s(X; \vartheta) \sim p_\vartheta \cdot L^1 \) for some \( \vartheta^* \in \Theta^* \subset \Theta \). We assume that
\[ E \left| \frac{p'_{\vartheta^*}(s(X; \vartheta))}{p_{\vartheta^*}(s(X; \vartheta))} s(X; \vartheta) \right| < \infty \quad \text{and/or} \quad E \left[ \left| \frac{p'_{\vartheta^*}(s(X; \vartheta))}{p_{\vartheta^*}(s(X; \vartheta))} \right| \right] < \infty. \]
By our characterizations, we have $s(X; \vartheta) \sim p_{\vartheta^*}L^1$ if, and only if, $T_{p_{\vartheta^*}}^{s(X; \vartheta)}(t) = F^{s(X; \vartheta)}(t)$ for every $t \in (L, R)$, where $F^{s(X; \vartheta)}$ denotes the distribution function of $s(X; \vartheta)$.

In order to test $H_0$, we consider a suitable measure of deviation between $T_{p_{\vartheta^*}}^{s(X; \vartheta)}$ and $F^{s(X; \vartheta)}$, that is

$$\delta\left(T_{p_{\vartheta^*}}^{s(X; \vartheta)}, F^{s(X; \vartheta)}\right), \quad (3.2)$$

where $\delta$ is a metric on a set containing both functions. This distance is zero under $H_0$, so large values of some suitable estimator of (3.2) will lead to the rejection of the hypothesis.

In the following, we propose a method to estimate (3.2). Therefore, let $X_1, \ldots, X_n$ be independent and identically distributed (iid.) copies of $X$. Setting $Y_{n,j} = s(X_j; \hat{\vartheta}_n)$, $j = 1, \ldots, n$, with a consistent estimator $\hat{\vartheta}_n = \hat{\vartheta}_n(X_1, \ldots, X_n)$ of $\vartheta \in \Theta$, we define an empirical version $\hat{T}_n^X$ of $T_{p_{\vartheta^*}}^{s(X; \vartheta)}$ by replacing the expectation through which $T_{p_{\vartheta^*}}^{s(X; \vartheta)}$ is defined with the arithmetic mean of the corresponding function in the expectation applied to $Y_{n,j}$, $j = 1, \ldots, n$. To that end, we use $\hat{\vartheta}_n^* = \hat{\vartheta}_n(Y_{n,1}, \ldots, Y_{n,n})$ as an estimator of $\vartheta^* \in \Theta^*$. Denoting by $\hat{F}_n$ the empirical distribution function of $Y_{n,1}, \ldots, Y_{n,n}$, we have as a suitable estimator of (3.2) the distance $\delta(\hat{T}_n^X, \hat{F}_n)$. Note that with additional thoughts on technical details the displayed approach also includes a handle for variable supports, i.e., if $\text{spt}(p_{\vartheta}) = [L_{\vartheta}, R_{\vartheta}]$ depends on the parameter $\vartheta$.

As has been witnessed by [4, 6, 7], tests of this type are noteworthy competitors to established tests. An advantage lies in the range of their applicability. A substantial proportion of known procedures relies on a comparison between theoretical moment generating functions, see [9, 18, 41], or characteristic functions, see [3, 14, 24], and their empirical pendants, or on a differential equation that characterizes the Laplace transformation, see [20, 22]. All of these share the unpleasant feature that in order to establish the theoretic basis for the test statistics, one has to know these transformations explicitly for the distribution in consideration. Since their calculation is not possible for every distribution, our suggestions provide a genuine alternative, since they require no more than the knowledge of the density function and its derivative. Moreover, our tests do not rely on a characterization that is tailored to one specific distribution. Instead, we provide a framework for testing fit to many different distributions, as indicated by the (not exhaustive) list of examples.
3.1 Tests for normality.

We show that the goodness-of-fit tests for normality proposed in [7] are included in our framework (cf. Example 2.8). To fix notation, we write \( p_\vartheta(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \), \( x \in \mathbb{R} \), where \( \vartheta = (\mu, \sigma^2) \in \Theta = \mathbb{R} \times (0, \infty) \), for the density function of a normal distribution with expectation \( \mu \) and variance \( \sigma^2 \). The function \( T^{X}_{\vartheta} = T^{X}_{p_\vartheta} \) from Theorem 2.5 takes the form

\[
T^{X}_{\vartheta}(t) = \mathbb{E} \left[ \frac{X-\mu}{\sigma^2} (X-t) \mathbbm{1}\{X \leq t\} \right], \quad t \in \mathbb{R},
\]

where \( X \) is a real-valued random variable with \( \mathbb{E}X^2 < \infty \). Taking into account the invariance under linear transformations of the class of normal distributions, the authors of [7] used the scaling function \( s(x; \vartheta) = (x-\mu)/\sigma \). Naturally, \( X \sim p_\vartheta \mathcal{L}^1 \) if, and only if, \( s(X; \vartheta) \sim p_\vartheta \mathcal{L}^1 \), where \( \vartheta^* = (0, 1) \), i.e., if \( s(X; \vartheta) \) follows the standard Gaussian law. Furthermore, we have

\[
\mathbb{E} \left[ \frac{|p'_{\vartheta^*}(s(X; \vartheta))|}{p_{\vartheta^*}(s(X; \vartheta))} \right] = \mathbb{E}|s(X; \vartheta)| \leq \frac{1}{\sigma} \left( \mathbb{E}|X| + |\mu| \right) < \infty
\]

and

\[
\mathbb{E} \left[ \frac{|p'_{\vartheta^*}(s(X; \vartheta))|}{p_{\vartheta^*}(s(X; \vartheta))} s(X; \vartheta) \right] = \mathbb{E}(s(X; \vartheta))^2 \leq \frac{1}{\sigma^2} \left( \mathbb{E}X^2 + 2|\mu|\mathbb{E}|X| + \mu^2 \right) < \infty.
\]

As a consequence, Theorem 2.5 states that \( X \) follows a normal distribution with parameter vector \( \vartheta = (\mu, \sigma^2) \) if, and only if, \( T^{s(X; \vartheta)}_{\vartheta} = F^{s(X; \vartheta)} \). For iid. copies \( X_1, \ldots, X_n \) of \( X \), we consider the sample mean \( \overline{X}_n = n^{-1} \sum_{j=1}^n X_j \) and variance \( S_n^2 = n^{-1} \sum_{j=1}^n (X_j - \overline{X}_n)^2 \) as consistent estimators of \( \mu \) and \( \sigma^2 \). We put

\[
Y_{n,j} = s(X_j; \overline{X}_n, S_n^2) = (X_j - \overline{X}_n)/S_n, \quad j = 1, \ldots, n,
\]

and notice that \( \hat{\vartheta}^*_n = \left( \overline{X}_n, S_n^2 \right) = (0, 1) \). Thus, an empirical version of \( T^{s(X; \vartheta)}_{\vartheta^*} \) is given by

\[
\hat{T}^{X}_n(t) = \frac{1}{n} \sum_{j=1}^n Y_{n,j}(Y_{n,j} - t) \mathbbm{1}\{Y_{n,j} \leq t\}, \quad t \in \mathbb{R}.
\]

It remains to compare \( \hat{T}^{X}_n \) with the empirical distribution function \( \hat{F}_n \) of \( Y_{n,1}, \ldots, Y_{n,n} \) by an appropriate measure of deviation. In particular, the authors of [7] considered \( \hat{T}^{X}_n \) and \( \hat{F}_n \) as random elements in the Hilbert space \( L^2(\mathbb{R}, \mathcal{B}^1, w(t) dt) \), where \( w \) is a suitable weight function, and chose as a metric the one induced by the Hilbert space norm. In accordance with our general considerations at the beginning of this section, the statistic has the form

\[
G^{(1)}_n = n \int_{\mathbb{R}} \left| \hat{T}^{X}_n(t) - \hat{F}_n(t) \right|^2 w(t) dt.
\]
Note that in [7] another statistic based on Corollary 2.24 and Example 2.25 is proposed, namely

\[ G_n^{(2)} = n \int_{\mathbb{R}} \left| \hat{T}_n^X(t) - \Phi(t) \right|^2 w(t) \, dt, \]

where \( \Phi \) is the distribution function of the standard normal distribution. Besides specifying weight functions for which the statistics have an explicit formula, the authors of [7] used the central limit theorem for random elements in separable Hilbert spaces to derive the limit distributions under the hypothesis \( H_0 \) in (3.1). Furthermore, they established the consistency of the test procedures against fixed alternatives with existing second moment and showed in a Monte Carlo simulation study that these tests are serious competitors to established procedures. The problem of testing for normality is still of interest in research, as evidenced by [18, 19, 21] and numerous preprints.

### 3.2 Tests for exponentiality.

The characterization of the exponential distribution via the mean residual life function is a special case of Theorem 2.10 (cf. Example 2.15). The authors of [4] used this characterization to construct associated tests in the sense described above. Denoting the density function of the exponential distribution with rate parameter \( \vartheta \in \Theta = (0, \infty) \) by \( p_{\vartheta}(x) = \vartheta \exp(-\vartheta x), \ x > 0 \), we have

\[ T_{\vartheta}^X = T_{p_{\vartheta}}^X = \mathbb{E}[\vartheta \min\{X, t\}], \ t > 0, \]

and \( T_{\vartheta}^X(t) = 0, \ t \leq 0 \). Here, \( X \) is a positive random variable with \( \mathbb{E}X < \infty \). To obtain a scale invariant test statistic, we consider the scaling function \( s(x; \vartheta) = \vartheta x \). Then \( X \) follows the exponential law with parameter \( \vartheta \) if, and only if, \( s(X; \vartheta) \sim p_{\vartheta^*} \mathcal{L}^1 \), where \( \vartheta^* = 1 \). We have

\[ \mathbb{E} \left| \frac{p_{\vartheta^*}(s(X; \vartheta))}{p_{\vartheta^*}(s(X; \vartheta))} s(X; \vartheta) \right| = \vartheta \mathbb{E}X < \infty, \]

and Theorem 2.10 implies that \( X \) has the exponential distribution with parameter \( \vartheta \) if, and only if, \( T_{\vartheta^*}^{s(X; \vartheta)} = F_{\vartheta^*}(s(X; \vartheta)) \). Let \( X_1, \ldots, X_n \) be iid. copies of \( X \) and denote by \( \overline{X}_n \) the sample mean, as a consistent estimator of \( \vartheta^{-1} \). Setting

\[ Y_{n,j} = s(X_j; \overline{X}_n^{-1}) = X_j / \overline{X}_n, \ j = 1, \ldots, n, \]

the empirical version of \( T_{\vartheta^*}^{s(X; \vartheta)} \) takes the form

\[ \hat{T}_n^X(t) = \frac{1}{n} \sum_{j=1}^n \min\{Y_{n,j}, t\}, \ t > 0. \]
With the uniform or a weighted $L^2$-norm, the test statistics become

$$L_n = \sqrt{n} \sup_{t > 0} \left| \widehat{F}_n^X(t) - \widehat{F}_n(t) \right| \quad \text{and} \quad G_n = n \int_0^\infty \left| \widehat{F}_n^X(t) - \widehat{F}_n(t) \right|^2 e^{-t} \, dt.$$ 

These are the statistics proposed and investigated in [4]. They showed that the limit distribution under the hypothesis coincides with the limiting null distribution of the classical Kolmogorov-Smirnov- and Cramér-von Mises statistic, respectively, when testing for uniformity over the unit interval. Furthermore, they proved the consistency of these test procedures against any fixed alternative distribution. The tests have already been included in the extensive comparative simulation study conducted by [1]. Adding a tuning parameter to the weight function of $G_n$ leads to the test statistic proposed in [5]. The recent papers [25, 30, 39, 40] show that tests for exponentiality are still of importance to the research community.

### 3.3 Tests for the Gamma distribution.

In their paper [6], the authors established the characterization result of Theorem 2.10 for the special case of the Gamma distribution and examined the corresponding goodness-of-fit statistics. Denote by $p_{\vartheta}(x) = \frac{\lambda^k}{\Gamma(k)} x^{k-1} e^{-x/\lambda}$, $x > 0$, where $\vartheta = (k, \lambda) \in \Omega = (0, \infty)^2$, the density function of the Gamma distribution with shape parameter $k$ and scale parameter $\lambda$.

The function $T_{\vartheta}^X = T_{\vartheta}^X$ from Theorem 2.10 takes the form

$$T_{\vartheta}^X(t) = \mathbb{E} \left[ \left( -\frac{k - 1}{X} + \frac{1}{\lambda} \right) \min\{X, t\} \right], \quad t > 0,$$

and $T_{\vartheta}^X(t) = 0$, $t \leq 0$. Here, $X$ is a positive random variable with $\mathbb{E}X < \infty$. To reflect the scale invariance of the class of Gamma distributions, choose the scaling function $s(x; \vartheta) = x/\lambda$. Apparently, we have $X \sim p_{\vartheta} \mathcal{L}^1$ if, and only if, $s(X; \vartheta) \sim p_{\vartheta^*} \mathcal{L}^1$, where $\vartheta^* = (k, 1) \in (0, \infty) \times \{1\} = \Theta^*$, and

$$\mathbb{E} \left| \frac{p'_{\vartheta^*}(s(X; \vartheta))}{p_{\vartheta^*}(s(X; \vartheta))} s(X; \vartheta) \right| \leq |k - 1| + \lambda^{-1} \mathbb{E}X < \infty.$$

By Theorem 2.10, $X$ follows a Gamma law with parameter vector $\vartheta = (k, \lambda)$ if, and only if, $T_{\vartheta^*}^{s(X; \vartheta)} = F_{\vartheta^*}^{s(X; \vartheta)}$. To construct the goodness-of-fit test, let $X_1, \ldots, X_n$ be iid. copies of $X$ and consider a consistent, scale equivariant estimator $\hat{\lambda}_n = \hat{\lambda}_n(X_1, \ldots, X_n)$ of $\lambda$ as well as a consistent, scale invariant estimator $\tilde{\lambda}_n = \tilde{\lambda}_n(X_1, \ldots, X_n)$ of $k$. We set

$$Y_{n,j} = s(X_j; \tilde{\lambda}_n, \hat{\lambda}_n) = X_j/\hat{\lambda}_n, \quad j = 1, \ldots, n.$$
Naturally, $\hat{\lambda}_n^* = \hat{\lambda}_n(Y_{n,1}, \ldots, Y_{n,n}) = 1$ and $\hat{k}_n^* = \hat{k}_n(Y_{n,1}, \ldots, Y_{n,n}) = \hat{k}_n(X_1, \ldots, X_n) = \hat{k}_n$ are consistent estimators of $\lambda^* = 1$ and $k^* = k$. An empirical version of $T^{s(X; \vartheta)}_{\vartheta^*}$ is given through

$$\hat{T}_n^X(t) = \frac{1}{n} \sum_{j=1}^{n} \left( -\hat{k}_n - \frac{1}{Y_{n,j}} + 1 \right) \min\{Y_{n,j}, t\}, \quad t > 0.$$ 

Similar to the test in Section 3.1, the authors of [6] considered the functions $\hat{T}_n^X$ and $\hat{F}_n = n^{-1} \sum_{j=1}^{n} \mathbb{1}\{Y_{n,j} \leq \cdot\}$ as random elements in the Hilbert space $L^2((0, \infty), \mathcal{B}_{>0}, w(t) \, dt)$, where $w$ is an appropriate weight function. They obtained the statistic

$$G_n = n \int_{0}^{\infty} \left\| \hat{T}_n^X(t) - \hat{F}_n(t) \right\|^2 w(t) \, dt,$$

derived the limit distribution under the hypothesis using the Hilbert space central limit theorem, and gave a proof of the consistency of this test procedure against fixed alternatives with existing expectation. Moreover, they explained how to implement the test using a parametric bootstrap and showed in a Monte Carlo simulation study that the test excels classical procedures and keeps up with the best Gamma tests proposed so far. Contributions like [22, 33, 38] indicate that testing fit to the Gamma distribution is also a topic of ongoing research.

### 3.4 Classical Procedures.

We consider the uniform distribution on the unit interval, $p(t) = \mathbb{1}_{(0,1)}(t), \ t \in \mathbb{R}$. Theorem 2.19 reads as follows (cf. Example 2.22). A random variable $X$ with values in the unit interval, and with distribution function $F$, is uniformly distributed if, and only if,

$$F(t) = T_p^X(t) = t \cdot \mathbb{1}_{(0,1)}(t) + \mathbb{1}_{[1,\infty)}(t)$$

for every $t \in \mathbb{R}$. Noting that $\hat{T}_n^X = T_p^X$ itself serves as an ‘empirical’ version, we obtain the statistics

$$K_n = \sqrt{n} \sup_{0 < t < 1} \left| \hat{F}_n(t) - T_p^X(t) \right| \quad \text{and} \quad \omega_n^2 = n \int_{0}^{1} \left| \hat{F}_n(t) - T_p^X(t) \right|^2 w(t) \, dt \quad \text{(3.3)}$$

for testing the uniformity hypothesis. Here, $w$ denotes some weight function and $\hat{F}_n$ is the empirical distribution function of $X_1, \ldots, X_n$ (which are iid. copies of $X$). Since $T_p^X$ is the distribution function of the uniform distribution, we recovered the classical characterization of a distribution via its distribution function. This applies to any probability measure on the real
line, and statistics of the form \((3.3)\) can be used to test for any fixed continuous distribution function \(F_0\). Assume that based on a sample \(X_1,\ldots,X_n\), where \(X_1\) has distribution function \(F\), we want to test the hypothesis \(F = F_0\) against the general alternative \(F \neq F_0\). Natural and well-understood test procedures are based on the statistics
\[
\widetilde{K}_n = \sqrt{n} \sup_{t \in \mathbb{R}} \left| \frac{\hat{F}_n(t) - F_0(t)}{w(F_0(t))} \right| 
\]
and
\[
\widetilde{\omega}_{n}^2 = n \int_{\mathbb{R}} \left| \frac{\hat{F}_n(t) - F_0(t)}{w(F_0(t))} \right|^2 w(F_0(t)) \, dF_0(t).
\]
Setting \(w \equiv 1\), these are the Kolmogorov-Smirnov- and Cramér-von Mises statistics, and using \(w(t) = (t(1-t))^{-1}\) in the integral statistic, we obtain the one from Anderson and Darling. They all share the pleasant feature that their distribution under the hypothesis does not depend on \(F_0\) and can be examined for \(F_0\) as the uniform distribution on the interval \((0,1)\). This essentially reduces them to the statistics in \((3.3)\), which are included in our framework.

For an account of the historical development of these classical procedures, a synoptic derivation of their limit distribution and an explanation on how to extend these tests to situations where the null hypothesis includes a whole (parametric) family of distributions, as well as for further references, we recommend [12].

4 Conclusions.

We devoted this work to the derivation of explicit characterizations for a large class of continuous univariate probability distributions. Our motivation was the fact that the characterization of the standard normal distribution as the unique fixed point of the zero-bias transformation reduces to an explicit formula for the distribution function of the transformed distribution. We extrapolated this formula to other distributions by applying the Stein type identity commonly used within the density approach. Research related to our characterizations concerns the study of distributional transformations, see [16] and [13]. While these are constructed from scratch and are used to prove Stein type characterizations, we took such a Stein identity for granted and dropped the ambition to obtain distributional transformations. Thus, starting with more information and demanding less structure from the transformations, we established better accessible explicit characterization formulae. In the last section, we discussed an immediate application. We illustrated how to use the characterizations for the construction of goodness-of-fit tests. The corresponding procedures for the normal, the exponential and the Gamma distribution have already been investigated in the literature, and they show very promising performance. The great advantage of our approach lies in the wide range of its applicability.
References

[1] J. S. Allison, L. Santana, N. Smit, and I. J. H. Visagie. An ‘apples to apples’ comparison of various tests for exponentiality. Computational Statistics, 32:1241–1283, 2017.

[2] A. D. Barbour. Stein’s method for diffusion approximations. Probability Theory and Related Fields, 84:297–322, 1990.

[3] L. Baringhaus and N. Henze. A consistent test for multivariate normality based on the empirical characteristic function. Metrika, 35:339–348, 1988.

[4] L. Baringhaus and N. Henze. Tests of fit for exponentiality based on a characterization via the mean residual life function. Statistical Papers, 41:225–236, 2000.

[5] L. Baringhaus and N. Henze. A new weighted integral goodness-of-fit statistic for exponentiality. Statistics & Probability Letters, 78:1006–1016, 2008.

[6] S. Betsch and B. Ebner. A new characterization of the Gamma distribution and associated goodness of fit tests. ArXiv e-prints, 1806.06028v2, 2018.

[7] S. Betsch and B. Ebner. Testing normality using a fixed point property of distributional transformations in the Stein characterization. ArXiv e-prints, 1803.07069, 2018.

[8] P. Billingsley. Probability and Measure. John Wiley and Sons, third edition, 1995.

[9] A. Cabaña and A. Quiroz. Using the empirical moment generating function in testing for the Weibull and the type I extreme value distributions. TEST, 14:417–432, 2005.

[10] S. Chatterjee and Q.-M. Shao. Nonnormal approximation by Steins method of exchangeable pairs with application to the Curie–Weiss model. The Annals of Applied Probability, 21:464–483, 2011.

[11] L. H. Y. Chen, L. Goldstein, and Q.-M. Shao. Normal approximation by Steins method. Springer, Berlin, 2011.

[12] E. del Barrio, J. A. Cuesta-Albertos, C. Matrán, S. Csörgö, C. M. Cuadras, T. de Wet, E. Giné, R. Lockhart, A. Munk, and W. Stute. Contributions of empirical and quantile processes to the asymptotic theory of goodness-of-fit tests. TEST, 9:1–96, 2000.
[13] C. Döbler. Distributional transformations without orthogonality relations. *Journal of Theoretical Probability*, 30:85–116, 2017.

[14] T. W. Epps and L. B. Pulley. A test for normality based on the empirical characteristic function. *Biometrika*, 70, 1983.

[15] L. Goldstein and G. Reinert. Stein’s method and the zero bias transformation with application to simple random sampling. *The Annals of Applied Probability*, 7:935–952, 1997.

[16] L. Goldstein and G. Reinert. Distributional transformations, orthogonal polynomials, and Stein characterizations. *Journal of Theoretical Probability*, 18:237–260, 2005.

[17] F. Götze. On the rate of convergence in the multivariate CLT. *The Annals of Probability*, 19:724–739, 1991.

[18] N. Henze and M. D. Jiménez-Gamero. A new class of tests for multinormality with i.i.d. and GARCH data based on the empirical moment generating function. *TEST*, 2018.

[19] N. Henze, M. D. Jiménez-Gamero, and S. G. Meintanis. Characterizations of multinormality and corresponding tests of fit, including for GARCH models. *Econometric Theory*, pages 1–37, 2018.

[20] N. Henze and B. Klar. Goodness-of-fit tests for the inverse Gaussian distribution based on the empirical Laplace transform. *Annals of the Institute of Statistical Mathematics*, 54:425–444, 2002.

[21] N. Henze and S. Koch. On a test of normality based on the empirical moment generating function. *Statistical Papers*, 2017.

[22] N. Henze, S. G. Meintanis, and B. Ebner. Goodness-of-fit tests for the Gamma distribution based on the empirical Laplace transform. *Communications in Statistics - Theory and Methods*, 41:1543–1556, 2012.

[23] H. M. Hudson. A natural identity for exponential families with applications in multiparameter estimation. *The Annals of Statistics*, 6:473–484, 1978.
[24] M. D. Jiménez-Gamero, V. Alba-Fernández, J. Muñoz-García, and Y. Chalco-Cano. Goodness-of-fit tests based on empirical characteristic functions. *Computational Statistics & Data Analysis*, 53:3957–3971, 2009.

[25] M. Jovanović, B. Milošević, Y. Y. Nikitin, M. Obradović, and K. Y. Volkova. Tests of exponentiality based on Arnold–Villasenor characterization and their efficiencies. *Computational Statistics & Data Analysis*, 90:100–113, 2015.

[26] C. Ley, G. Reinert, and Y. Swan. Steins method for comparison of univariate distributions. *Probability Surveys*, 14:1–52, 2017.

[27] C. Ley and Y. Swan. Stein’s density approach and information inequalities. *Electronic Communications in Probability*, 18, 2013.

[28] Y. V. Linnik. Linear forms and statistical criteria i, ii. *Selected Translations in Mathematical Statistics and Probability*, 3:1–40, 41–90. Originally published 1953 in the Ukrainian Mathematical Journal, Vol. 5, pp. 207–243, 247–290 (in Russian), 1962.

[29] Y. Y. Nikitin. Tests based on characterizations, and their efficiencies: A survey. *Acta et Commentationes Universitatis Tartuensis de Mathematica*, 21:3–24, 2017.

[30] H. A. Noughabi. Testing exponentiality based on the likelihood ratio and power comparison. *Annals of Data Science*, 2:195–204, 2015.

[31] F. J. O’Reilly and M. A. Stephens. Characterizations and goodness of fit tests. *Journal of the Royal Statistical Society. Series B (Methodological)*, 44:353–360, 1982.

[32] E. A. Peköz and A. Röllin. New rates for exponential approximation and the theorems of Rényi and Yaglom. *The Annals of Probability*, 39:587–608, 2011.

[33] B. Plubin and P. Siripanich. An alternative goodness-of-fit test for a Gamma distribution based on the independence property. *Chiang Mai Journal of Science*, 44:1180–1190, 2017.

[34] B. L. S. Prakasa Rao. Characterizations of distributions through some identities. *Journal of Applied Probability*, 16:903–909, 1979.

[35] N. Ross. Fundamentals of Steins method. *Probability Surveys*, 8:210–293, 2011.
[36] C. Stein. Approximate computation of expectations. *Lecture Notes - Monograph Series*, 7, 1986.

[37] C. Stein, P. Diaconis, S. Holmes, and G. Reinert. *Use of exchangeable pairs in the analysis of simulations*, volume 46 of *Lecture Notes – Monograph Series*. Institute of Mathematical Statistics, 2004.

[38] J. A. Villaseñor and E. González-Estrada. A variance ratio test of fit for Gamma distributions. *Statistics & Probability Letters*, 96:281–286, 2015.

[39] K. Y. Volkova and Y. Y. Nikitin. Exponentiality tests based on Ahsanullah’s characterization and their efficiency. *Journal of Mathematical Sciences*, 204:42–54, 2015.

[40] V. Zardasht, S. Parsi, and M. Mousazadeh. On empirical cumulative residual entropy and a goodness-of-fit test for exponentiality. *Statistical Papers*, 56:677–688, 2015.

[41] A. A. Zghoul. A goodness of fit test for normality based on the empirical moment generating function. *Communications in Statistics – Simulation and Computation*, 39:1292–1304, 2010.

S. Betsch and B. Ebner,
Institute of Stochastics,
Karlsruhe Institute of Technology (KIT),
Englerstr. 2, D-76133 Karlsruhe.
E-mail: Steffen.Betsch@student.kit.edu and Bruno.Ebner@kit.edu