Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 53 (5) (2024), 1453 – 1483
DOI : 10.15672/hujms.1482377

Research Article

On some mixtures of the Kies distribution

Tsvetelin Stefanov Zaevski*[1,2], Nikolay Kyurkchiev[1,3]

[1] *Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Acad. Georgi Bonchev Str., Block 8, Sofia, 1113, Bulgaria*
[2] *Faculty of Mathematics and Informatics, Sofia University St. Kliment Ohridski, 5, James Bourchier Blvd., Sofia, 1164, Bulgaria*
[3] *Faculty of Mathematics and Informatics, University of Plovdiv Paisii Hilendarski, 236, Bulgaria Blvd., Plovdiv, 4027, Bulgaria*

Abstract

The purpose of this paper is to explore some mixtures, discrete and continuous, based on the Kies distribution. Some conditions for convergence are established. We study the probabilistic properties of these mixtures. Special attention is taken to the so-called Hausdorff saturation. Several models are examined in detail – bimodal, multimodal, and mixtures based on binomial, geometric, exponential, gamma, and beta distributions. We provide some numerical experiments for real-life tasks – one for the Standard and Poor's 500 financial index and another for unemployment insurance issues. In addition, we check the consistency of the proposed estimator using generated data of different sizes.

Mathematics Subject Classification (2020). 41A40, 41A46, 60E05, 62E17

Keywords. Probability mixtures, Compound distributions, Kies distribution, Exponential distribution, Weibull distribution, Hausdorff saturation

1. Introduction

The exponential style distributions are widely used in the stochastic theory as well as in many real-life fields. The outstanding theoretical importance of the original exponential distribution arises from the fact that it is the unique continuous one that exhibits memorylessness.† Later Weibull introduces a power dependence in [45]. Thus the flexibility of these models is significantly increased. The price of this is the loss of the memorylessness. Nevertheless, they are some of the most applicable distributions in survival analysis. This implies the wide use in different areas such as engineering sciences, meteorology and hydrology, communications and telecommunications, chemical and metallurgical industry, medicine and epidemiology, insurance and banking, etc – we refer to [8,13,25,31,40,46,48]. Note that these distributions are stated on the infinite domain $(0, \infty)$. Alternatively,

*Corresponding Author.
Email addresses: t_s_zaevski@math.bas.bg, t_s_zaevski@abv.bg (T. S. Zaevski), nkyurk@math.bas.bg (N. Kyurkchiev)
Received: 11.05.2024; Accepted: 10.09.2024

†For example, this property allows defining the very applicable class of the Lévy processes – stochastic processes with independent, identical distributed increments.

Kies[19] applies the fractional-linear transform $t = \frac{y}{y+1} \Leftrightarrow y = \frac{t}{1-t}$ to the Weibull distribution and thus changes the domain into the interval $(0,1)$. If one uses the transform $t = \frac{by+a}{y+1} \Leftrightarrow y = \frac{t-a}{b-t}$, $a < b$, instead of $t = \frac{y}{y+1}$, then the distribution will be stated on the interval (a,b) – see for example [27,28,50]. Several extensions of this distribution are available in the science literature. In [3,5,29,30] the authors propose a power transformation to define new families. Some composite distributions are constructed in [51] – see also [6,7]. The distributional properties of the minimum and maximum of several Kies distributions are discussed in [52]. Some trigonometric transformations are proposed in [53]. Thus, all these modifications fall into the large class of exponential style distributions and thus they inherit the large applicability of their predecessors.

On the other hand, mixing probability distributions is a powerful method for enlarging their flexibility and applicability due to several inherent features. First, this approach can combine the characteristics of two, three, or more distributions – this way non-homogeneous data can be explored. Second, many new bi- and multi-modal models that exhibit some desired features can be constructed. Third, the substitution of a distribution's parameter with a random variable can strengthen some characteristics but keep the main properties of the original model. Also, the approach of mixing can be applied to the stochastic processes, not only to the random variables. The subordination is an example of this in some sense.

The usefulness of this method is supported by the large number of sources devoted to the topic. The applications of such kind distributions can be found in many real-life areas, such as sociology [24], food industry [26], information and communication technologies [20,33,49,54], engineering [15], nanotechnologies [44], biology [21], meteorology [37], genetics [17], medicine [22,34,39,43], finance [10], economy and energy industry [9,12,47], etc. In addition, many new mixture models appear recently – exponential [36], based on the distributions of Shanker [2], Topp-Leone [38], log-Bilal [23], Akash [35], Lindley [33], Kumaraswamy [18].

In the present paper, we build new probability distributions mixing a Kies-style family and assuming that its parameters are random variables. We specify the resulting distribution via its cumulative function defining it as the average of the original ones. We establish some necessary conditions which keep the main characteristics of the initial Kies family. The probabilistic properties of the new distributions are explored. Several special cases are examined in detail – discrete mixtures (bi- and multi-modal, binomial, geometric) as well as mixtures based on continuous distributions – exponential, gamma, and beta. Several reasons motivate this choice. First, the bi- and multi-modal mixtures are very intuitive and thus they are often applied to incorporate the behavior of the underlying distributions. Note that the multi-modal models are prone to over-fitting and thus they have to be used very carefully and only when there is strong evidence that the studied object exhibits namely such behavior. The rest discrete mixtures we provide (binomial and geometric) illustrate how their probability structures influence the resulting distribution. Furthermore, the geometric distribution is the unique discrete memorylessness one. Next, we examine its continuous analogue – the exponential mixture – as well as its generalization – the gamma one. In addition, another related model is explored based on the beta distribution. This way we can check the practical implementation of these upgradeable models as well as the impact of the driving parameters. An interesting example that arises is the fact that the standard uniform distribution can be viewed as an exponential Kies mixture.

Another important task we discuss is the so-called saturation, defined as the Hausdorff distance between the cumulative distribution function and a Γ-shaped curve that connects its endpoints. The so-established term can be viewed as a measure of the speed of occurrence or as an indicator for a critical point. We derive a semi-closed form formula in the general case and apply it to the above-mentioned particular mixtures. In addition, this formula turns to explicit for the exponential mixtures. For some additional studies devoted to the Hausdorf saturation, we refer to [41, 51–53].

We apply the derived results to statistical samples generated from two real-life areas – the financial industry and the social sphere. The first example is about the calm and volatile periods for the Standard and Poor's 500 (S&P 500) index, the second one is about the unemployment insurance issues. These statistical data exhibit quite different behavior. The density of the first one seems to have an infinitely large initial value, whereas the density of the second one has zero endpoints and one peak. Different kinds of parameters of the Kies distribution and the resulting mixtures can approximate both behaviors. We calibrate the above-mentioned mixtures (bimodal, multimodal, binomial, geometric, exponential, gamma, and beta). The derived results are discussed in detail – we have to mention that the exponential distribution and its gamma extension produce very realistic results.

Last but not least, we investigate the asymptotic behavior of the proposed estimator. We do this by generating random samples with different sizes (1 000, 10 000, 100 000, and 1 000 000) based on the exponential mixed Kies distribution. It is chosen due to the low number of parameters that define it – only two. This way we can ignore the possible error coming from the multidimensional optimization. Our tests strongly confirm the consistency of the estimating approach.

The paper is structured as follows. We present the base we use later in Section 2. The Kies mixtures are defined and examined in Section 3. The Hausdorff saturation is investigated in Section 4. Some particular examples are considered in Section 5. The applications of the Kies mixtures are discussed in 6. We check the proposed estimator through generated samples in Section 7.

2. Preliminaries

We shall use the following notations: a large letter for the cumulative distribution function (CDF) of a distribution, the over-lined letter for the complementary cumulative distribution function (CCDF), the corresponding small letter for the probability density function (PDF), and the letter ψ for the moment generating function (MGF). Thus if $F(t)$ is the CDF, then $\overline{F}(t)$, $f(t)$, and $\psi(t)$ are the corresponding CCDF, PDF, and MGF, respectively. We shall preserve the common notation F for the mixed distribution defined in Section 3 whereas the letter H will be used for the underlying Kies one.

The Kies distribution on the domain $(0,1)$ is defined via its CDF, CCDF, or PDF:

$$H(t) = 1 - \exp\left(-\lambda\left(\frac{t}{1-t}\right)^{\beta}\right) \tag{2.1}$$

$$\overline{H}(t) = \exp\left(-\lambda\left(\frac{t}{1-t}\right)^{\beta}\right) \tag{2.2}$$

$$h(t) = \lambda\beta\frac{t^{\beta-1}}{(1-t)^{\beta+1}}\exp\left(-\lambda\left(\frac{t}{1-t}\right)^{\beta}\right). \tag{2.3}$$

$$\tag{2.4}$$

The shape of the probability density function is obtained in proposition 2.1 from [51]:

Proposition 2.1. *The value of the PDF at the right endpoint of the domain is zero,* $h(1) = 0$*. Let the function* $\alpha(t)$ *for* $t \in (0,1)$ *be defined as*

$$\alpha(t) := \lambda\beta\left(\frac{t}{1-t}\right)^{\beta} - (2t + \beta - 1). \tag{2.5}$$

The following statements for PDF (2.3) *related to the position of the power* β *with respect to the unit number hold.*

(1) *If* $\beta > 1$*, then PDF* (2.3) *is zero in the left domain's endpoint,* $h(0) = 0$*. Function* (2.5) *has a unique root for* $t \in (0,1)$*, we denote it by* t_2*. The PDF increases for* $t \in (0, t_2)$ *having a maximum for* $t = t_2$ *and decreases for* $t \in (t_2, 1)$*.*

(2) *If* $\beta = 1$*, then the left limit of the PDF is* $h(0) = \lambda$*. If* $\lambda \geq 2$*, then the PDF is a decreasing from* λ *to* 0 *function. Otherwise, if* $\lambda < 2$*, then we need the value* $t_2 = 1 - \frac{\lambda}{2}$ *– note that* $t_2 \in (0,1)$*. The PDF starts from the value* λ *for* $t = 0$*, increases to a maximum for* $t = t_2$*, and decreases to zero.*

(3) *If* $\beta < 1$*, then* $h(0) = \infty$*. The derivative of function* (2.5) *is*

$$\alpha'(t) = \lambda\beta^2 \frac{t^{\beta-1}}{(1-t)^{\beta+1}} - 2. \tag{2.6}$$

Let $\bar{t}$ *be defined as* $\bar{t} := \frac{1-\beta}{2}$*. The PDF is a decreasing function when* $\alpha'(\bar{t}) \geq 0$*.*

Suppose that $\alpha'(\bar{t}) < 0$*. In this case, derivative* (2.6) *has two roots in the interval* $(0,1)$ *– we denote them by* $\bar{t}_1$ *and* $\bar{t}_2$*. If* $\alpha(\bar{t}_2) \geq 0$*, then the PDF decreases in the whole distribution domain. Otherwise, if* $\alpha(\bar{t}_2) < 0$*, then function* (2.5) *has two roots in the interval* $(0,1)$ *too – we notate them by* t_1 *and* t_2*. The PDF starts from infinity, decreases in the interval* $(0, t_1)$ *having a local minimum for* $t = t_1$*, increases for* $t \in (t_1, t_2)$ *having a local maximum for* $t = t_2$*, and decreases to zero for* $t \in (t_2, 1)$*.*

Generally said, Proposition 2.1 shows that the PDF may exhibit several forms. In all cases, the right endpoint is zero. Also, it may start from infinity and decrease to zero or have one local minimum and another local maximum. Second, it may start from a finite point and decrease to zero or first to increase to a peak and then to decrease to zero. Finally, the PDF may start from zero having a bell form.

3. Kies mixtures

We shall define now the main objects of our study – the mixtures based on the Kies distribution. We do this assuming that its parameters are random variables and averaging with respect to their laws. We need to impose some conditions to guarantee that the resulting distribution inherits the properties of its predecessors. An important characteristic of the original Kies distribution is that the initial point of its PDF can be zero, finite, or infinite valued – see proposition 2.1. We prove that the mixtures preserve this feature and obtain the conditions that distinguish the different cases. Furthermore, the main distributional properties are investigated. It is shown at the end of this section how the mixtures can be defined at a joint probability space.

We need the following lemma to define the Kies mixtures.

Lemma 3.1. *Let* λ *and* β *be positive random variables on some probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and* $H(t; \lambda, \beta)$ *be the CDFs of a Kies distributed family. Then the function* $F(\cdot)$*, defined as*

$$F(t) = \mathbb{E}\left[H(t; \lambda, \beta)\right], \tag{3.1}$$

is continuous and increases from zero to one.

Proof. Continuity follows from the dominated convergence theorem since $0 \leq H(t; \lambda, \beta) \leq 1$ for all sample events. As a consequence $F(0) = 0$ and $F(1) = 1$. Function (3.1) is increasing since all functions $H(t; \lambda, \beta)$ increase. □

Remark 3.2. Although the symbols λ and β are often associated with population parameters in the statistical literature, in Lemma 3 we use them for random variables. We do this to keep the symbols of Section 2 for the original Kies distribution as well as some other traditional notations. For example, the symbol λ is usually used for the intensity of the exponential distribution – a predecessor of the Kies one.

Based on Lemma 3.1 we can define the mixed distribution in the following way:

Definition 3.3. Let us impose the following conditions on the random variables λ and β:

$$\mathbb{E}\left[\frac{\beta}{\lambda^{\frac{1}{\beta}}}\left(\frac{\beta+1}{\beta}\right)^{\frac{\beta+1}{\beta}}\right]<\infty \tag{3.2}$$

$$\mathbb{E}\left[\lambda\right]<\infty \tag{3.3}$$

$$\mathbb{E}\left[\lambda\beta\right]<\infty. \tag{3.4}$$

$$\tag{3.5}$$

The mixture distribution is defined by its CDF, $F(\cdot)$, through formula (3.1). Its parameters are defined via the random variables λ and β.

The importance of these conditions will be seen later when we discuss some particular examples. Note that they are sufficient but not necessary and thus one may impose other requirements that lead to similar results.

We need the following lemma before establishing the result for the PDF of the mixture.

Lemma 3.4. *Let a and b be positive constants. The function*

$$g(x)=x^{b+1}e^{-ax^b} \tag{3.6}$$

achieves its maximum at the positive real half-line for $x=\left(\frac{b+1}{ab}\right)^{\frac{1}{b}}$ and it is

$$\left(\frac{b+1}{eab}\right)^{\frac{b+1}{b}}. \tag{3.7}$$

Proof. The proof follows the presentation of the derivative of function (3.6)

$$g'(x)=x^be^{-ax^b}\left(b+1-abx^b\right). \tag{3.8}$$

□

The next two propositions are for the PDFs of the mixed distributions. The second one establishes the behavior near zero.

Proposition 3.5. *Let $t\in(0,1]$. The following statements for the CCDF and PDF of the mixture hold:*

$$\overline{F}(t)=\mathbb{E}\left[\overline{H}(t;\lambda,\beta)\right] \tag{3.9}$$

$$f(t)=\mathbb{E}\left[h(t;\lambda,\beta)\right]. \tag{3.10}$$

$$\tag{3.11}$$

Proof. The equality $\overline{F}(t)=\mathbb{E}\left[\overline{H}(t;\lambda,\beta)\right]$ is obvious. Next, we prove the statement for the PDF. We can write

$$\begin{aligned}F'(t)&=\lim_{\epsilon\to0}\frac{\mathbb{E}\left[H(t+\epsilon;\lambda,\beta)\right]-\mathbb{E}\left[H(t;\lambda,\beta)\right]}{\epsilon}\\&=\lim_{\epsilon\to0}\mathbb{E}\left[\frac{H(t+\epsilon;\lambda,\beta)-H(t;\lambda,\beta)}{\epsilon}\right]\\&=\lim_{\epsilon\to0}\mathbb{E}\left[h(\tau(\epsilon);\lambda,\beta)\right]\end{aligned} \tag{3.12}$$

for some $\tau(\epsilon) \in (t, t+\epsilon)$ due to the mean value theorem. Using Lemma 3.4 for $a = \lambda t^{\beta}$, $b = \beta$, and $x = \frac{1}{1-t}$ we obtain

$$\begin{aligned}\mathbb{E}\left[|h(t;\lambda,\beta)|\right] &= \mathbb{E}\left[\lambda\beta\frac{t^{\beta-1}}{(1-t)^{\beta+1}}\exp\left(-\lambda\left(\frac{t}{1-t}\right)^{\beta}\right)\right] \\ &\leq \mathbb{E}\left[\lambda\beta t^{\beta-1}\left(\frac{\beta+1}{e\lambda\beta t^{\beta}}\right)^{\frac{\beta+1}{\beta}}\right] \\ &\leq \mathbb{E}\left[\frac{\beta\exp\left(-\frac{\beta+1}{\beta}\right)}{t^2\lambda^{\frac{1}{\beta}}}\left(\frac{\beta+1}{\beta}\right)^{\frac{\beta+1}{\beta}}\right].\end{aligned} \tag{3.13}$$

Hence $\mathbb{E}\left[|h(t;\lambda,\beta)|\right] < \infty$ in a neighborhood of every point from the interval $t \in (0,1)$ due to condition (3.2). We can apply now the dominated convergence theorem to obtain the desired result (3.10) for the PDF. □

Obviously, the shape of the mixture PDF is closely related to the PDFs of the Kies family as well as to the behavior of the random variables λ and β. However, results in a closed form cannot be obtained in the general case. Nonetheless, we can prove the following proposition that characterizes the endpoints of the PDF.

Proposition 3.6. *The right endpoint of the PDF is $f(1) = 0$. The left endpoint can be derived via the following alternatives. Note that we shall use the symbol $\mathbb{E}$ for the expectation with respect to the measure $\mathbb{P}$.*

(1) *If $\mathbb{P}(\beta > 1) = 1$, then $f(0) = 0$.*
(2) *If $\mathbb{P}(\beta = 1) > 0$ and $\mathbb{P}(\beta < 1) = 0$, then $f(0) = \mathbb{Q}(\beta = 1)\mathbb{E}[\lambda]$. The probability measure $\mathbb{Q}$ is equivalent to $\mathbb{P}$ with Radon-Nikodym derivative*

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{\lambda}{\mathbb{E}[\lambda]}. \tag{3.14}$$

Note that the measure $\mathbb{Q}$ really exists since the random variable λ is positive and $\mathbb{E}[\lambda] < \infty$ due to condition (3.3).
(3) *If $\mathbb{P}(\beta < 1) > 0$, then $f(0) = \infty$.*

Proof. The value $f(1) = 0$ can be derived via inequality (3.13) and the dominated convergence theorem. Let us turn to the left endpoint. We can rewrite formula (3.10) as

$$f(t) = \mathbb{E}\left[h(t;\lambda,\beta)I_{\beta<1}\right] + \mathbb{E}\left[h(t;\lambda,\beta)I_{\beta=1}\right] + \mathbb{E}\left[h(t;\lambda,\beta)I_{\beta>1}\right]. \tag{3.15}$$

Suppose first that $\mathbb{P}(\beta > 1) = 1$. Let us define the measure $\mathbb{L}$ for a fixed t via the Radon-Nikodym derivative

$$\frac{d\mathbb{L}}{d\mathbb{P}} = \frac{h(t;\lambda,\beta)}{\mathbb{E}\left[h(t;\lambda,\beta)\right]}. \tag{3.16}$$

Note that for every $t \in (0,1]$ the expectation $\mathbb{E}\left[h(t;\lambda,\beta)\right]$ is finite due to inequality (3.13). Hence the sum of the first and second expectations from formula (3.15) can be obtained as

$$\mathbb{E}\left[h(t;\lambda,\beta)I_{\beta\leq 1}\right] = \mathbb{E}\left[h(t;\lambda,\beta)\right]\mathbb{L}(\beta \leq 1) = 0, \tag{3.17}$$

because the measures $\mathbb{L}$ and $\mathbb{P}$ are equivalent and $\mathbb{P}(\beta \leq 1) = 0$. Inequality (3.4) allows us to take the limit $t \to 0$ in the third expectation from formula (3.15) and using the first statement of proposition 2.1 to derive $f(0) = 0$.

Suppose now that $\mathbb{P}(\beta = 1) > 0$ and $\mathbb{P}(\beta < 1) = 0$. Analogously as above, we can derive that the values of the first and third expectations from formula (3.15) are zero. We derive the desired result by changing the measure from $\mathbb{P}$ to $\mathbb{Q}$ in the second expectation.

It left to consider the case $\mathbb{P}(\beta < 1) > 0$. The third statement of proposition 2.1 shows that $h(t; \lambda(\omega), \beta(\omega))$ tends to infinity for $t \to 0$ for every sample event ω such that $\beta(\omega) < 1$. Let $t \in (0, 1]$ and M be a positive constant. We define the set $\Omega_{t,M}$ as

$$\Omega_{t,M} = \{\omega \in \Omega : (\beta < 1) \,\&\, (h(t; \lambda, \beta) > M \;\forall\; u < t)\}. \tag{3.18}$$

Hence,

$$\mathbb{E}\left[h(t; \lambda, \beta) I_{\beta<1}\right] \geq M\mathbb{P}(\Omega_{t,M}). \tag{3.19}$$

Having in mind that $\lim_{t\to 0} \Omega_{t,M} = \{\omega : \; \beta < 1\}$, we see that (3.19) leads to

$$\lim_{t\to 0} \mathbb{E}\left[h(t; \lambda, \beta) I_{\beta<1}\right] \geq M\mathbb{P}(\beta < 1). \tag{3.20}$$

We conclude that the third expectation of formula (3.15) tends to infinity because inequality (3.20) holds for all constants M and $\mathbb{P}(\beta < 1) > 0$. This finishes the proof. □

Next, we provide a useful result which states that the CCDF can be derived through the moment-generating function of λ if β is non-random.

Corollary 3.7. *Suppose that β is a constant. The CCDF of a Kies mixture can be derived as the MGF of the random variable λ taken at the point $-\left(\frac{t}{1-t}\right)^{\beta}$*

$$\overline{F}(t) = \psi_{\lambda}\left(-\left(\frac{t}{1-t}\right)^{\beta}\right). \tag{3.21}$$

As a consequence, the PDF turns into

$$f(t) = \beta\frac{t^{\beta-1}}{(1-t)^{\beta+1}}\psi'_{\lambda}\left(-\left(\frac{t}{1-t}\right)^{\beta}\right). \tag{3.22}$$

Proof. The result is true due to formulas (2.2) and (3.9). □

Finally, we discuss how the triple (ξ, λ, β) consisting of the mixed distribution ξ and the random variables λ and β can be defined at a joint probability space.

Remark 3.8. We can define the mixed Kies distribution jointly with the pair (λ, β) on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let the measure $\mu(dx_1, dx_2)$ on $\mathbb{R}^+ \times \mathbb{R}^+$ be associated with the random variables (λ, β) and ξ be the related Kies mixture. We define the triple (ξ, λ, β) via the joint distribution

$$\mathbb{P}(\xi \in A, \lambda \in B_1, \beta \in B_2) = \int_{t\in A}\int_{x\in B} h(t, x_1, x_2)\,\mu(dx_1, dx_2)dt \tag{3.23}$$

for arbitrary subsets A and B of the sets $(0, 1)$ and $\mathbb{R}^+ \times \mathbb{R}^+$. We can easily check that formula (3.1) holds:

$$
\begin{aligned}
F(t) &= \mathbb{P}(\xi < t) \\
&= \int_0^t \int_{x\in\{\mathbb{R}^+\times\mathbb{R}^+\}} h(t, x_1, x_2)\, \mu(dx_1, dx_2) dt \\
&= \int_{x\in\{\mathbb{R}^+\times\mathbb{R}^+\}} \left(\int_0^t h(t, x_1, x_2) dt \right) \mu(dx_1, dx_2) \\
&= \mathbb{E}\left[H(t; \lambda, \beta)\right].
\end{aligned}
\tag{3.24}
$$

We have used above Fubini's theorem to interchange the order of integration. Note that the defined in that way random variable ξ is independent of the pair (λ, β) only when λ and β are constants.

4. Saturation in the Hausdorff sense

The Hausdorff distance can be defined in the sense of [32]:

Definition 4.1. Let us consider the max-norm in $\mathbb{R}^2$: if A and B are the points $A = (t_A, x_A)$ and $B = (t_B, x_B)$, then $\|A - B\| := \max\{|t_A - t_B|, |x_A - x_B|\}$. The Hausdorff distance $d(g, h)$ between two curves g and h in $\mathbb{R}^2$ is

$$
d(g, h) := \max\left\{ \sup_{A\in g} \inf_{B\in h} \|A - B\|, \sup_{B\in h} \inf_{A\in g} \|A - B\| \right\}. \tag{4.1}
$$

The Hausdorff distance can be viewed as the highest optimal path between the curves. Next, we define the saturation of a distribution:

Definition 4.2. Let $F(\cdot)$ be the CDF of a distribution stated on the interval $(0, 1)$. Its saturation is the Hausdorff distance between the completed graph of $F(\cdot)$ and the curve consisting of two lines – one vertical between the points $(0, 0)$ and $(0, 1)$ and another horizontal between the points $(0, 1)$ and $(1, 1)$.

The following corollary for the saturation holds.

Corollary 4.3. *The saturation d of a mixed-Kies random variable is the unique solution of the equation*

$$
F(d) + d = 1. \tag{4.2}
$$

Proof. Equation (4.2) is true due to Definitions 4.1 and 4.2. Its root is unique because $l(t) = F(t) + t - 1$ is an increasing continuous function with endpoints $l(0) = -1 < 0$ and $l(1) = 1 > 0$. □

The next two theorems provide a semi-closed form formula for the saturation.

Theorem 4.4. *Let the positive random variable τ be such that*

$$
\lambda = \tau \left(\frac{1}{\mathbb{E}\left[e^{-\tau}\right]} - 1 \right)^{\beta}. \tag{4.3}
$$

The saturation of the Kies mixture can be derived as

$$
d = \mathbb{E}\left[e^{-\tau}\right]. \tag{4.4}
$$

Note that $0 < \mathbb{E}\left[e^{-\tau}\right] < 1$.

Proof. Suppose that formula (4.3) holds. Using equations (2.1) and (3.1) and having in mind $0 < \mathbb{E}[e^{-\tau}] < 1$ we obtain

$$\begin{aligned} 1 - F(\mathbb{E}[e^{-\tau}]) &= \mathbb{E}\left[1 - H(\mathbb{E}[e^{-\tau}];\lambda,\beta)\right] \\ &= \mathbb{E}\left[\exp\left\{-\lambda\left(\frac{\mathbb{E}[e^{-\tau}]}{1-\mathbb{E}[e^{-\tau}]}\right)^{\beta}\right\}\right] \\ &= \mathbb{E}\left[\exp\left\{-\tau\left(\frac{1}{\mathbb{E}[e^{-\tau}]}-1\right)^{\beta}\left(\frac{\mathbb{E}[e^{-\tau}]}{1-\mathbb{E}[e^{-\tau}]}\right)^{\beta}\right\}\right] \\ &= \mathbb{E}[e^{-\tau}]. \end{aligned} \tag{4.5}$$

We have to use Corollary 4.3 to finish the proof. □

Theorem 4.5. *The random variable τ from Theorem 4.4 that satisfies condition* (4.3) *exists and it is unique.*

Proof. Suppose that there exist two random variables τ_1 and τ_2 satisfying equation (4.3). Hence,

$$\tau_1\left(\frac{1}{\mathbb{E}[e^{-\tau_1}]}-1\right)^{\beta} = \tau_2\left(\frac{1}{\mathbb{E}[e^{-\tau_2}]}-1\right)^{\beta}, \tag{4.6}$$

which can be rewritten as

$$\left(\frac{\tau_1}{\tau_2}\right)^{\frac{1}{\beta}} = \frac{\mathbb{E}[e^{-\tau_1}]}{1-\mathbb{E}[e^{-\tau_1}]}\frac{1-\mathbb{E}[e^{-\tau_2}]}{\mathbb{E}[e^{-\tau_2}]}. \tag{4.7}$$

We can see that the right-hand side of equation (4.7) is a deterministic constant. Therefore, the ratio $\frac{\tau_1}{\tau_2}$ can be presented as

$$\frac{\tau_1}{\tau_2} = c^{\beta} \tag{4.8}$$

for some constant c. Combining equations (4.7) and (4.8) we derive

$$\begin{aligned} c &= \frac{\mathbb{E}\left[e^{-c^{\beta}\tau_2}\right]}{1-\mathbb{E}\left[e^{-c^{\beta}\tau_2}\right]}\frac{1-\mathbb{E}[e^{-\tau_2}]}{\mathbb{E}[e^{-\tau_2}]} \\ &= \left(\frac{1}{1-\mathbb{E}\left[e^{-c^{\beta}\tau_2}\right]}-1\right)\frac{1-\mathbb{E}[e^{-\tau_2}]}{\mathbb{E}[e^{-\tau_2}]}. \end{aligned} \tag{4.9}$$

The right-hand side of equation (4.9) is a decreasing function with respect to variable c. Hence, if we consider (4.9) as an equation with respect to c, then it has at most one solution. We can easily check that this root is $c = 1$, which means $\tau_1 = \tau_2$.

We turn to the existence task. Let the function $\gamma(x)$ be defined as

$$\gamma(x) = x\left[\frac{1}{\mathbb{E}\left[e^{-\lambda x^{\beta}}\right]}-1\right]. \tag{4.10}$$

Note that this function increases from zero to infinity in the interval $x \in [0,\infty)$. Hence, the equation $\gamma(x) = 1$ has a unique solution – we denote it by $\overline{x}$. We shall show that the random variable

$$\tau = \lambda\overline{x}^{\beta} \tag{4.11}$$

satisfies equation (4.3). Using the equality $\gamma(\overline{x}) = 1$, we derive

$$\left(\frac{\lambda}{\tau}\right)^{\frac{1}{\beta}} = \frac{1}{\overline{x}} = \frac{1 - \mathbb{E}\left[e^{-\lambda \overline{x}^{\beta}}\right]}{\mathbb{E}\left[e^{-\lambda \overline{x}^{\beta}}\right]} = \frac{1 - \mathbb{E}\left[e^{-\tau}\right]}{\mathbb{E}\left[e^{-\tau}\right]}, \tag{4.12}$$

which is equivalent to equation (4.3). This finishes the proof. □

The following theorem is an immediate corollary of Theorems 4.4 and 4.5 and gives an approach for deriving the saturation of the mixed-Kies distributions.

Theorem 4.6. *Let $\overline{x}$ be the solution of the equation $\gamma(x) = 1$, where the function $\gamma(x)$ is defined by formula* (4.10)*. Note that this root is unique in the interval $x \in [0, \infty)$. Then the saturation can be obtained via formula* (4.4)*, where τ is defined by equation* (4.11)*. Combining equations $\gamma(x) = 1$,* (4.4)*, and* (4.11)*, we see that the saturation can be derived also as*

$$d = \frac{\overline{x}}{\overline{x} + 1}. \tag{4.13}$$

Based on Theorem 4.6, we can establish the following algorithm for deriving the saturation.

We can obtain the saturation through the next three steps:

(1) We derive $\overline{x}$ as the solution of $\gamma(x) = 1$, where the function $\gamma(x)$ is given by formula (4.10).
(2) We obtain τ via equation (4.11).
(3) We derive the saturation through formula (4.4) or equivalently through (4.13).

5. Some examples

We can devise the set of all mixtures introduced by Definition 3.3 into two main classes – discrete and continuous. We shall consider separately several examples of both kinds. The numerical simulations are prepared via MATLAB. The corresponding codes are provided at https://github.com/zhuszhus/Tsvetelin-Zaevski/tree/main/.

5.1. Discrete mixtures

Suppose that we have $n \leq \infty$ original Kies-distributions – if $n = \infty$ we want the set of these distributions to be countable. Note that conditions (3.2)-(3.4) are satisfied when $n < \infty$. The possible values of the random variables λ and β shall be denoted by λ_i and β_i, $i = 1, 2, ..., n$. The probabilities of these values are denoted by $p_i > 0$. We can write the PDF, CDF, and CCDF as

$$\begin{aligned} f(t) &= \sum_{i=1}^{n} p_i h(t; \lambda_i, \beta_i) \\ F(t) &= \sum_{i=1}^{n} p_i H(t; \lambda_i, \beta_i) \\ \overline{F}(t) &= \sum_{i=1}^{n} p_i \overline{H}(t; \lambda_i, \beta_i), \end{aligned} \tag{5.1}$$

where the functions $h(t; \lambda_i, \beta_i)$, $H(t; \lambda_i, \beta_i)$, and $\overline{H}(t; \lambda_i, \beta_i)$ are given by equations (2.1)-(2.3). Proposition 3.6 leads to the following results:

Corollary 5.1. *Let the set $\{1, 2, ..., n\}$ be devised into the subsets A_1, A_2, and A_3 such that (i) if $i \in A_1$, then $\beta_i < 1$; (ii) if $i \in A_2$, then $\beta_i = 1$; (iii) if $i \in A_3$, then $\beta_i > 1$. The following statements hold:*

(1) *If $A_1 \equiv A_2 \equiv \emptyset$, then $f(0) = 0$.*
(2) *If $A_1 \equiv \emptyset$ but $A_2 \neq \emptyset$, then $f(t) = \sum_{i \in A_2} p_i \lambda_i$.*
(3) *If $A_1 \neq \emptyset$, then $f(0) = \infty$.*

5.1.1. Bimodal distribution. Let the parameter β be deterministic and larger than one, say $\beta = 2$. This guarantees that the initial point of the PDF is zero due to Corollary 5.1. We assume that λ is a random variable and achieves two values $\lambda \in \{0.1, 2\}$ with probabilities $p_1 = p_2 = 0.5$ or $p_1 = 0.25$ and $p_2 = 0.75$. Proposition 2.1 shows that the PDFs of the original Kies distributions are zero in the domain's endpoints and have a unique maximum. The mixture distribution exhibits bi-modality – all PDFs can be seen in Figure 0a. Of course, some values may lead to a unimodal distributions – for example, the random variable $\lambda \in \{1, 2\}$, together with the same rest parameters, leads to such PDF, see Figure 0b.

Another example for bi-modality can be seen in Figure 0c – the assumed parameters are $\lambda \in \{2, 0.5\}$ and $\beta \in \{2, 1\}$. The probabilities are again $p_1 = p_2 = 0.5$ or $p_1 = 0.25$ and $p_2 = 0.75$. The main difference with the previous examples is that one of the β-values is one and another is larger than one. Corollary 5.1, the second statement, explains why the left endpoint of the PDF is larger than zero but finite. Something more, its value is $p_2\lambda_2 - \frac{1}{4}$ or $\frac{3}{8}$ for both possibilities for p_1 and p_2.

Finally, we present a bimodal distribution with an infinite left endpoint – see Figure 0d. The chosen parameters are $\lambda \in \{2, 0.5\}$ and $\beta \in \{0.2, 2\}$; the probabilities are the same as above. The third statement of Corollary 5.1 confirms that the mixture's left endpoints are the infinity since $\beta_1 < 1$.

Let us discuss the Hausdorff saturation defined in Section 4. We shall follow Algorithm 4. We first have to derive the solution of equation $\gamma(x) = 1$, which now turns into

$$x \left(\frac{1}{\sum_{j=1}^{n} p_j e^{-\lambda_j x^{\beta_j}}} - 1 \right) = 1. \tag{5.2}$$

Estimating $n = 2$, $\beta = 2$, $p_1 = p_2 = 0.5$, $\lambda_1 = 0.1$, and $\lambda_2 = 2$, we derive $\overline{x} = 1.0338$. Hence, the values of $\tau = \lambda \overline{x}^{\beta}$ are $\tau_1 = 0.1069$ and $\tau_2 = 2.1374$ and the saturation is $d = 0.5083$. We can easily check that equation (4.2) holds.

For the second case ($p_1 = 0.25$, $p_2 = 0.75$), we derive $\overline{x} = 0.7986$, $\tau_1 = 0.0638$, $\tau_2 = 1.2755$, and $d = 0.4440$. The CDF together with the saturation can be viewed in Figure 2a. The saturation is indicated by a red circle. Note that the red lines form a square with vertices $(0, 1-d)$, $(d, 1-d)$, $(d, 1)$, and $(0, 1)$.

5.1.2. Multimodal distributions. We present a multimodal Kies mixture based on four original Kies distributions in Figure 0e. The used parameters are $\lambda \in \{0.1, 0.5, 5, 10\}$ and $\beta = 2$. The probabilities are assumed to be equal, i.e. $p_1 = p_2 = p_3 = p_4 = 0.25$. The first statement of Corollary 5.1 shows that the mixture left endpoint is zero since $\beta > 1$.

Applying Algorithm 4 and having in mind equation (5.2), we derive $\overline{x} = 0.7700$ and $\tau \in \{0.0593, 0.2965, 2.9646, 5.9291\}$. Thus the saturation is $d = 0.4350$.

5.1.3. Binomial underlying distribution. Let the random variable $\lambda - 1$ be binomial distributed with parameters (n, p) and β be a constant. Note that λ is positive. We have for the probabilities p_i, $i = 1, 2, ..., n+1$:

$$p_i = \mathbb{P}(\lambda = i) = \binom{n}{i-1} p^{i-1} (1-p)^{n+1-i}. \tag{5.3}$$

We turn to deriving the CCDF. We cannot use directly Corollary 3.7 because we know the distribution of the random variable $\lambda - 1$, not λ. However, we have

$$\begin{aligned} \overline{F}(t) &= \sum_{i=1}^{n+1} p_i \overline{H}(t; \lambda_i, \beta) \\ &= \sum_{i=1}^{n+1} p_i \exp\left(-i\left(\frac{t}{1-t}\right)^{\beta}\right) \\ &= \exp\left(-\left(\frac{t}{1-t}\right)^{\beta}\right) \sum_{i=0}^{n} p_i \exp\left(-(i-1)\left(\frac{t}{1-t}\right)^{\beta}\right). \end{aligned} \tag{5.4}$$

We can recognize in the sum above the MGF of the binomial distribution applied to the term $-\left(\frac{t}{1-t}\right)^{\beta}$. Hence, if we denote this function by $\psi(x)$,

$$\psi(x) = (1 - p + pe^{x})^{n}, \tag{5.5}$$

then we derive

$$\begin{aligned} \overline{F}(t) &= \exp\left(-\left(\frac{t}{1-t}\right)^{\beta}\right) \psi\left(-\left(\frac{t}{1-t}\right)^{\beta}\right) \\ &= \exp\left(-\left(\frac{t}{1-t}\right)^{\beta}\right) \left(1 - p + p\exp\left(-\left(\frac{t}{1-t}\right)^{\beta}\right)\right)^{n}. \end{aligned} \tag{5.6}$$

Having in mind that the derivative of the function $\left(\frac{t}{1-t}\right)$ is $\frac{1}{(1-t)^2}$, we obtain for the binomial Kies mixture:

$$\begin{aligned} f(t) &= \beta \frac{t^{\beta-1}}{(1-t)^{\beta+1}} \exp\left(-\left(\frac{t}{1-t}\right)^{\beta}\right) \left(1 - p + p\exp\left(-\left(\frac{t}{1-t}\right)^{\beta}\right)\right)^{n} \\ &+ np\beta \frac{t^{\beta-1}}{(1-t)^{\beta+1}} \exp\left(-2\left(\frac{t}{1-t}\right)^{\beta}\right) \left(1 - p + p\exp\left(-\left(\frac{t}{1-t}\right)^{\beta}\right)\right)^{n-1} \\ &= \beta \frac{t^{\beta-1}}{(1-t)^{\beta+1}} \exp\left(-\left(\frac{t}{1-t}\right)^{\beta}\right) \left(1 - p + p\exp\left(-\left(\frac{t}{1-t}\right)^{\beta}\right)\right)^{n-1} \times \\ &\times \left(1 - p + p(n+1)\exp\left(-\left(\frac{t}{1-t}\right)^{\beta}\right)\right). \end{aligned} \tag{5.7}$$

Combining equations (5.6) and (5.7), we derive for the hazard function

$$\Lambda(t) := \frac{f(t)}{\overline{F}(t)} = \beta \frac{t^{\beta-1}}{(1-t)^{\beta+1}} \frac{1 - p + p(n+1)\exp\left(-\left(\frac{t}{1-t}\right)^{\beta}\right)}{1 - p + p\exp\left(-\left(\frac{t}{1-t}\right)^{\beta}\right)}. \tag{5.8}$$

We present in Figure 0f the PDFs of the Kies mixtures when $\beta = 2$, $n \in \{10, 50\}$, and $p \in \{0.5, 0.25\}$. Corollary 5.1 shows that the initial value of the PDF is zero.

We use again Algorithm 4 to obtain the Hausdorff saturation d. Remind that $\lambda = 1, 2, ..., 11$ when $n = 10$. In the case $p = 0.25$, we derive for the values of p_i, $\overline{x}$, and τ:

$$\begin{aligned}
&p \in \{0.0563, 0.1877, 0.2816, 0.2503, 0.1460, 0.0584, 0.0162, 0.0031, 0.0004,\\
&2.8610 \times 10^{-5}, 9.5367 \times 10^{-7}\}\\
&\overline{x} = 0.5632 \qquad (5.9)\\
&\tau \in \{0.3172, 0.6344, 0.9515, 1.2687, 1.5859, 1.9031, 2.2203, 2.5374,\\
&2.8546, 3.1718, 3.4890\}.
\end{aligned}$$

We obtain the saturation as $d = 0.3603$ via formula (4.13). Alternatively, we can view the expectation in formula (4.10) as the MGF, i.e.

$$\mathbb{E}\left[e^{-\lambda x^{\beta}}\right] = e^{-x^2}\left(1 - p + pe^{-x^2}\right)^n \tag{5.10}$$

Remind that λ is one plus a binomial distributed random variable. Thus, the equation $\gamma(x) = 1$ turns into

$$x\left[\frac{e^{x^2}}{\left(1 - p + pe^{-x^2}\right)^n} - 1\right] = 1. \tag{5.11}$$

Following the same approach we derive for the saturation: $\{\overline{x} = 0.4510, d = 0.3108\}$ when $\{\beta = 2, n = 10, p = 0.5\}$; $\{\overline{x} = 0.3279, d = 0.2470\}$ when $\{\beta = 2, n = 50, p = 0.25\}$; and $\{\overline{x} = 0.2506, d = 0.2004\}$ when $\{\beta = 2, n = 50, p = 0.5\}$. The CDF together with the Hausdorff saturation are presented in Figure 2b.

5.1.4. Geometric underlying distribution. We investigate now a Kies mixture based on a geometrically distributed random variable λ with parameter p on the support $\{1, 2, ...\}$. This is a mixture between an infinite number of original Kies distributions. Note that conditions (3.2)-(3.3) are satisfied because $\mathbb{E}\left[\lambda^{-\frac{1}{\beta}}\right] < 1$. The parameter β is again assumed to be deterministic. The probabilities p_i for $i = 1, 2, ...$ are

$$p_i = \mathbb{P}(\lambda = i) = p(1 - p)^{i-1}. \tag{5.12}$$

We can obtain the mixture CCDF through Corollary 3.7 as the MGF of the geometric distributions $\psi(x)$ taken at the point $-\left(\frac{t}{1-t}\right)^{\beta}$. Having in mind that this function is

$$\psi(x) = \frac{pe^x}{1 - (1 - p)e^x}, \tag{5.13}$$

we conclude

$$\overline{F}(t) = \frac{p}{\exp\left(\left(\frac{t}{1-t}\right)^{\beta}\right) + p - 1}. \tag{5.14}$$

Differentiating, we derive the PDF as

$$f(t) = \frac{p\beta \exp\left(\left(\frac{t}{1-t}\right)^{\beta}\right) t^{\beta-1}}{\left(\exp\left(\left(\frac{t}{1-t}\right)^{\beta}\right) + p - 1\right)^2 (1 - t)^{\beta+1}}. \tag{5.15}$$

Combining equations (5.14) and (5.15), we derive for the hazard function

$$\Lambda(t) = \frac{\beta \exp\left(\left(\frac{t}{1-t}\right)^{\beta}\right) t^{\beta-1}}{\left(\exp\left(\left(\frac{t}{1-t}\right)^{\beta}\right) + p - 1\right)(1 - t)^{\beta+1}}. \tag{5.16}$$

We present in Figure 1a the PDFs of the obtained mixtures considering the parameter p in the set $\{0.25, 0.5, 0.75\}$. We assume also that $\beta = 2$ and thus the initial point of the PDF is zero due to Corollary 5.1 . The saturations can be derived using Algorithm 4. We again view the expectation in function $\gamma(\cdot)$, given by (4.10), as MGF (5.13) taken at the point $-x^{\beta}$. Thus, equation $\gamma(x) = 1$ turns into

$$x\left(e^{x^{\beta}} - 1\right) = p. \tag{5.17}$$

Solving this equation we derive for its root $\overline{x}$ and the related saturation:
$\{\overline{x} = 0.5931, d = 0.3723\}$ when $\{\beta = 2, p = 0.25\}$; $\{\overline{x} = 0.7245, d = 0.4201\}$ when $\{\beta = 2, p = 0.5\}$; and $\{\overline{x} = 0.8097, d = 0.4474\}$ when $\{\beta = 2, p = 0.75\}$. We present the CDF in the case $\{\beta = 2, p = 0.25\}$ in Figure 2c – there can be seen the saturation as a red point too.

5.2. Continuous distributions

We consider now several mixtures based on the assumption that the parameter λ follows a continuous distribution supported on the positive real half-line. The random variable β is assumed to be deterministic.

5.2.1. Exponential distribution. Let $\beta > 1$, which means that the initial point of the PDF of the mixture is zero. We shall see later that conditions (3.2)-(3.4) hold only for such values of β. We assume that the random variable λ is exponentially distributed with intensity θ. Thus, its PDF is $p(x) = \theta e^{-\theta x}$. The MGF is defined for $x < \theta$ and it is

$$\psi(x) = \frac{\theta}{\theta - x}. \tag{5.18}$$

Applying Corollary 3.7 we see that the CCDF of the Kies mixture can be obtained after the substitution $x = -\left(\frac{t}{1-t}\right)^{\beta}$. Note that the condition $x < \theta$ is satisfied. Hence,

$$\overline{F}(t) = \frac{\theta(1-t)^{\beta}}{\theta(1-t)^{\beta} + t^{\beta}}. \tag{5.19}$$

Differentiating, we derive the PDF of the exponential mixture

$$f(t) = \frac{\theta\beta t^{\beta-1}(1-t)^{\beta-1}}{\left(\theta(1-t)^{\beta} + t^{\beta}\right)^2}. \tag{5.20}$$

Combining equations (5.19) and (5.20), we derive for the hazard function

$$\Lambda(t) = \frac{\beta t^{\beta-1}}{(1-t)\left(\theta(1-t)^{\beta} + t^{\beta}\right)}. \tag{5.21}$$

Several PDFs can be seen in Figure 1b – the intensity parameter θ is amongst $\{0.5, 1, 2, 5\}$. We find the saturation using Algorithm 4 having in mind that if β is a fixed constant, then the expectation in formula (4.10) for the function $\gamma(\cdot)$ is MGF (5.18) evaluated at the point $-x^{\beta}$. In the case of the exponential distribution, the equation $\gamma(x) = 1$ can be solved explicitly, and it leads to

$$\overline{x} = \theta^{\frac{1}{\beta+1}} \Leftrightarrow d = \frac{\theta^{\frac{1}{\beta+1}}}{\theta^{\frac{1}{\beta+1}} + 1}. \tag{5.22}$$

Thus, we derive $\{\overline{x} = 0.7937, d = 0.4425\}$ when $\{\beta = 2, \theta = 0.5\}$; $\{\overline{x} = 1, d = 0.5\}$ when $\{\beta = 2, \theta = 1\}$; $\{\overline{x} = 1.2599, d = 0.5575\}$ when $\{\beta = 2, \theta = 2\}$; $\{\overline{x} = 1.7100, d = 0.6310\}$

when $\{\beta = 2, \theta = 5\}$. The CDF, together with the saturation, in the case $\{\beta = 2, \theta = 1\}$ can be seen in Figure 2d.

Let us check conditions (3.2)-(3.4). Obviously, requirements (3.3) and (3.4) hold. We need to consider the expectation $\mathbb{E}\left[\lambda^{-\frac{1}{\beta}}\right]$ for condition (3.2). It can be written as

$$\mathbb{E}\left[\lambda^{-\frac{1}{\beta}}\right] = \int_0^\infty x^{-\frac{1}{\beta}} \theta e^{-\theta x} dx = \theta^{\frac{1}{\beta}} \int_0^\infty y^{\left(1-\frac{1}{\beta}\right)-1} e^{-y} dy. \tag{5.23}$$

The last integral converges only when $\beta > 1$ – the limit is the gamma function $\Gamma\left(1 - \frac{1}{\beta}\right)$. On the opposite, if $\beta \leq 1$, integral (5.23) diverges, and hence condition (3.2) is not satisfied.

Let us consider now the case $\beta = 1$. Function (3.1) again defines a distribution. Its CCDF can be derived once again through the MGF and thus it and the PDF turn into

$$\begin{aligned} \overline{F}(t) &= \frac{\theta(1-t)}{\theta(1-t)+t} \\ f(t) &= \frac{\theta}{(\theta(1-t)+t)^2}. \end{aligned} \tag{5.24}$$

We can see that Propositions 3.5 and 3.6 do not hold. For example, $\mathbb{E}[h(1, \lambda, \beta)] = 0$, but $f(1) = \theta$. We can formulate also the following result

Proposition 5.2. *The uniform distribution on the interval* $(0, 1)$ *can be viewed as a Kies mixture with* $\beta = 1$ *and exponentially distributed* λ *with intensity one.*

Proof. We recognize the CCDF and PDF of the uniform distribution in formulas (5.24) when $\theta = 1$. □

5.2.2. Gamma distribution. Suppose now that the random variable λ is gamma distributed with the shape and rate parameters α and θ, respectively. Note that this distribution generalizes the exponential one. The gamma PDF is

$$p(x) = \frac{\theta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\theta x}. \tag{5.25}$$

Let β be again a constant. We shall check first when conditions (3.2)-(3.4) are satisfied. Obviously, we need to consider the expectation $\mathbb{E}\left[\lambda^{-\frac{1}{\beta}}\right]$. We have

$$\mathbb{E}\left[\lambda^{-\frac{1}{\beta}}\right] = \frac{\theta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-\frac{1}{\beta}-1} e^{-\theta x} dx = \theta^{\frac{1}{\beta}} \int_0^\infty y^{\alpha-\frac{1}{\beta}-1} e^{-y} dy. \tag{5.26}$$

The integral above converges only when $\beta > \frac{1}{\alpha}$ – its value is $\Gamma\left(\alpha - \frac{1}{\beta}\right)$. We have for the gamma MGF

$$\psi(x) = \left(\frac{\theta}{\theta - x}\right)^\alpha \tag{5.27}$$

for $x < \theta$. Applying Corollary 3.7, we obtain for the mixture CCDF

$$\overline{F}(t) = \psi\left(-\left(\frac{t}{1-t}\right)^\beta\right) = \left(\frac{\theta(1-t)^\beta}{\theta(1-t)^\beta + t^\beta}\right)^\alpha. \tag{5.28}$$

Differentiating, we derive the PDF as

$$f(t) = \frac{\alpha \theta^\alpha \beta t^{\beta-1} (1-t)^{\alpha\beta-1}}{\left(\theta(1-t)^\beta + t^\beta\right)^{\alpha+1}}. \tag{5.29}$$

Combining equations (5.28) and (5.29), we derive for the hazard function

$$\Lambda(t) = \frac{\alpha\beta t^{\beta-1}\left(\theta(1-t)^{\beta}+t^{\beta}\right)}{(1-t)\left(\theta(1-t)^{\beta}+t^{\beta}\right)}. \tag{5.30}$$

We present some PDFs in Figure 1c for $\alpha = 2$ and $\theta \in \{0.5, 1, 2, 5\}$. We again use Algorithm 4 to derive the saturation. The equation $\gamma(x) = 1$ now can be written as

$$x\left(\left(\theta + x^{\beta}\right)^{\alpha} - \theta^{\alpha}\right) = \theta^{\alpha}, \tag{5.31}$$

because the expectation in (4.10) is MGF (5.27) evaluated at the point $-\lambda x^{\beta}$.
Hence, $\{\overline{x} = 0.5731, d = 0.3643\}$ when $\{\alpha = 2, \beta = 2, \theta = 0.5\}$; $\{\overline{x} = 0.7332, d = 0.4230\}$ when $\{\alpha = 2, \beta = 2, \theta = 1\}$; $\{\overline{x} = 0.9361, d = 0.4835\}$ when $\{\alpha = 2, \beta = 2, \theta = 2\}$; and $\{\overline{x} = 1.2894, d = 0.5632\}$ when $\{\alpha = 2, \beta = 2, \theta = 5\}$.

Suppose now that $\beta = 1$. Formula (5.29) leads to $f(0) = \frac{\alpha}{\theta}$ in which we recognize the expectation of the gamma distribution, $\mathbb{E}[\lambda]$. This is in accordance with the second statement of Proposition 3.6. Note that $\mathbb{Q}(\beta = 1) = 1$. The PDFs for the same values of α and θ can be viewed in Figure 1d. Their initial values are $\frac{\alpha}{\theta}$, particularly 4, 2, 0.5, and 0.4. The saturations can be derived in the same way as above – they are $\{\overline{x} = 0.4196, d = 0.2956\}$ when $\{\alpha = 2, \beta = 1, \theta = 0.5\}$; $\{\overline{x} = 0.6180, d = 0.3820\}$ when $\{\alpha = 2, \beta = 1, \theta = 1\}$; $\{\overline{x} = 0.9032, d = 0.4746\}$ when $\{\alpha = 2, \beta = 1, \theta = 2\}$; and $\{\overline{x} = 1.476, d = 0.5961\}$ when $\{\alpha = 2, \beta = 1, \theta = 5\}$.

Finally, we present some PDFs of the third kind assuming that $\beta = 0.7$, see Figure 1e. Note that $f(0) = \infty$ due to the third statement of Proposition 3.6. The saturations can be obtained analogously – $\{\overline{x} = 0.3512, d = 0.2599\}$ when $\{\alpha = 2, \beta = 0.7, \theta = 0.5\}$; $\{\overline{x} = 0.5615, d = 0.3596\}$ when $\{\alpha = 2, \beta = 0.7, \theta = 1\}$; $\{\overline{x} = 0.8855, d = 0.4696\}$ when $\{\alpha = 2, \beta = 0.7, \theta = 2\}$; and $\{\overline{x} = 1.5885, d = 0.6137\}$ when $\{\alpha = 2, \beta = 0.7, \theta = 5\}$. The CDF in the case $\{\alpha = 2, \beta = 0.7, \theta = 0.5\}$ together with the related saturation can be seen in Figure 2e.

5.2.3. Beta distribution. Assume now that the random variable λ is beta distributed with parameters α and θ, i.e. its PDF and MGF are

$$\begin{aligned} p(x) &= \frac{x^{\alpha-1}(1-x)^{\theta-1}}{B(\alpha,\theta)} \\ \psi(x) &= {}_1F_1(\alpha, \alpha+\theta, x), \end{aligned} \tag{5.32}$$

where ${}_1F_1(\cdot,\cdot,\cdot)$ is the confluent hypergeometric function of the first kind and the beta function is defined as the ratio

$$B(\alpha,\theta) = \frac{\Gamma(\alpha)\Gamma(\theta)}{\Gamma(\alpha+\theta)}. \tag{5.33}$$

We shall check when condition (3.2) is satisfied. We have

$$\mathbb{E}\left[\lambda^{-\frac{1}{\beta}}\right] = \int_0^{\infty} x^{-\frac{1}{\beta}} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} dx = \frac{1}{B(\alpha,\beta)} \int_0^{\infty} x^{\alpha-\frac{1}{\beta}-1}(1-x)^{\beta-1} dx. \tag{5.34}$$

The integral above converges when $\beta > \frac{1}{\alpha}$. Applying Corollary 3.7, we obtain the CCDF of the Kies-beta-mixture

$$\overline{F}(t) = \psi\left(-\left(\frac{t}{1-t}\right)^{\beta}\right) = {}_1F_1\left(\alpha, \alpha+\theta, -\left(\frac{t}{1-t}\right)^{\beta}\right). \tag{5.35}$$

Differentiating in equation (5.35) and using the formula for the confluent hypergeometric function's derivative – see for example [14], page 1023, formula 9.213 – we derive for the PDF of the mixture

$$f(t) = \frac{\alpha\beta}{\alpha+\theta} {}_1F_1\left(\alpha+1, \alpha+\theta+1, -\left(\frac{t}{1-t}\right)^{\beta}\right) \frac{t^{\beta-1}}{(1-t)^{\beta+1}}. \tag{5.36}$$

The hazard function $\Lambda(t) := \frac{f(t)}{\overline{F}(t)}$ can be derived by combining equations (5.35) and (5.36). Let us check the behavior of PDF (5.36) when $t \to 1$ or equivalently $-\left(\frac{t}{1-t}\right)^{\beta} \to -\infty$. Having in mind that ${}_1F_1(\alpha+1, \alpha+\theta+1, -x)$ tends asymptotically to $\frac{\Gamma(\alpha+\theta+1)}{\Gamma(\theta)} x^{-(\alpha+1)}$ when $x \to \infty$ – see [1], page 508, formula 13.5.1 – we derive

$$\begin{aligned} f(1) &= \lim_{t\to 1} \frac{\alpha\beta}{\alpha+\theta} \frac{\Gamma(\alpha+\theta+1)}{\Gamma(\theta)} \left(\frac{t}{1-t}\right)^{-\beta(\alpha+1)} \frac{t^{\beta-1}}{(1-t)^{\beta+1}} \\ &= \alpha\beta \frac{\Gamma(\alpha+\theta)}{\Gamma(\theta)} \lim_{t\to 1} (1-t)^{\alpha\beta-1}. \end{aligned} \tag{5.37}$$

We conclude that $f(1) = \infty$ when $\beta < \frac{1}{\alpha}$; $f(1) = \alpha\beta\frac{\Gamma(\alpha+\theta)}{\Gamma(\theta)}$ when $\beta = \frac{1}{\alpha}$; and $f(1) = 0$ when $\beta > \frac{1}{\alpha}$. Thus, the right endpoint is zero if condition (3.2) is satisfied due to equation (5.34). Having in mind that the value of the confluent hypergeometric function in the zero is one, we find a confirmation of Proposition 3.6 in formula (5.36). Note that if $\beta = 1$, then $f(0) = \frac{\alpha}{\alpha+\theta}$, which is the expectation of the beta distribution. Some PDFs can be seen in Figure 1f – the considered parameters are $\alpha = 3$, $\theta = 1$, and $\beta \in \{0.5, 1, 2\}$. Note that the condition $\beta > \frac{1}{\alpha}$ holds. We derive the saturations through Algorithm 4. Now, the equation $\gamma(x) = 1$ can be written as

$$x\left[\frac{1}{{}_1F_1(\alpha, \alpha+\theta, -x^{\beta})} - 1\right] = 1. \tag{5.38}$$

The derived values are $\{\overline{x} = 0.9577, d = 0.4892\}$ when $\{\alpha = 3, \beta = 0.5, \theta = 1\}$; $\{\overline{x} = 0.9697, d = 0.4923\}$ when $\{\alpha = 3, \beta = 1, \theta = 1\}$; $\{\overline{x} = 0.9806, d = 0.4951\}$ when $\{\alpha = 3, \beta = 2, \theta = 1\}$. The CDF and the saturation for the last triple are presented in Figure 2f.

6. An application

We shall check now the benefits that the proposed mixture gives, using two empirical samples – one arising from the financial markets and another for the unemployment insurance issues. These data are considered in [51, 52] too. We use a modification of the classical least square errors approach. First, note that we need to scale the empirical data because the considered distributions are stated at the domain $(0, 1)$. Let the total number of observations be denoted by N. We devise the domain into m sub-intervals and denote by N_i the number of observations that fall in the i-th one, $i = 1, 2, ..., m$. We derive the empirical PDF values as $l_i^{\text{emp}} = \frac{mN_i}{N}$. We assign these values to the centers of the sub-intervals. We denote by $l_i^{\text{th}}(\gamma)$ the theoretical PDF values at these points for a Kies mixture with parameter's set γ. The cost function is defined as

$$L(\gamma) := \sum_{i=1}^{m} \left| \ln\left(l_i^{\text{emp}} + \epsilon\right) - \ln\left(l_i^{\text{th}}(\gamma) + \epsilon\right) \right|. \tag{6.1}$$

We want to obtain the parameters γ which minimize function (6.1). We introduce the logarithmic correction because some Kies distributions tend to infinity in the left endpoint. The constant ϵ is necessary because some empirical values can be equal to zero which will lead to the minus infinity for the logarithm. We set this constant to $\epsilon = 0.01$.

We shall compare several mixtures – original Kies (A1), bimodal (A2), multimodal (A3), binomial (A4), geometric (A5), exponential (A6), gamma (A7), and beta (A8). We assume for the last five mixtures that the random variable λ does not exhibit the corresponding distribution, but its linear transformation. This increases significantly the applicability of the proposed models. If the original random variable is denoted by ξ, then the MGF of the resulting one can be derived as $\psi_{a\xi+b}(x) = e^{bx}\psi_{\xi}(ax)$. The PDF of the random variable $\lambda = a\xi + b$ can be obtained via Corollary 3.7 – see also Section 5. All resulting PDFs are reported in Appendix A. Note that if $b > 0$, then condition (3.2) is satisfied because the term $(at + b)$ in the integral

$$\mathbb{E} = \left[\lambda^{-\frac{1}{\beta}}\right] = \int\limits_{\mathbb{R}} (at + b)^{-\frac{1}{\beta}} f(t)\, dt \tag{6.2}$$

does not influence the possible singularity of $f(t)$ in the zero. The case $b = 0$ is considered in Section 5.2.

There are several reasons to choose these models. First, the original Kies distribution can be viewed as a benchmark. Thus we can see whether the generalizations made lead to a large enough contribution in describing the real data. Second, the proposed models include the most important class of mixtures – bi- and multi-modal, discrete, and continuous ones. Also, since these distributions form an upgradable sequence, we can decide which one provides a large enough impact. Last but not least, the linear transformation applied to the mixing distribution significantly increases the model flexibility.

6.1. S&P500 index

The statistical sample consists of a total of 10 717 daily observations for the S&P500 index in the period between January 2, 1980, and July 01, 2022. We look for the market shocks defined as the dates at which the index falls with more than two percent. More precisely, we are interested in the length of the periods between two shocks measured in working days – the calm periods. Their number is 357 and they vary between 1 and 950. We divide them by 1 000 to fit the distribution's domain. For more details see [51] – there can be found also the original Kies calibration. The interval $(0, 1)$ is devised into $m = 50$ sub-intervals for the current test. All mixture estimations we prepare are reported in the first part of Table 1. The corresponding densities are presented in Figure 3a. The inner figure is for the distribution's core – the interval $(0.03, 0.2)$. We can see that although the statistical data suggests that the initial density value is infinity, some of the estimated distributions do not exhibit this feature. In terms of Proposition 3.6 and Corollary 5.1, this means that the variable β, random or not, is larger than one. This is the case for the three-modal estimation as well as for the binomial, geometric, exponential, and gamma mixtures.

We also present in Table 1 the expectations that the different models generate as well as the 95% confidence intervals. The empirical averaged value is 0.03, i.e. 30 days between market shocks (remind that the number of days is divided by 1 000). We can see that all models estimate the expectation conservatively in the sense that they produce lower values – between 22 days (binomial mixture) and 25.9 days (beta mixture). The confidence intervals have similar forms – they are stated on the left side of the distribution domain. The lower one is for the exponential mixture $(0, 0.0831)$ – between 0 and 83.1 days.

6.2. Unemployment insurance issues

The second example is related to the monthly observations of the unemployment insurance issues for the period between 1971 and 2018 – a total of 574 observations in the range $[49\,263, 308\,352]$. The data can be found at https://data.worlddatany-govns8zxewg or in [42], pp. 162-164. Some statistical experiments based on the same data are provided also in [4,16,52,55]. We need first to process the statistical sample to make it convenient for our investigation. In [52] the same statistical sample is divided by 50 000 and thus the new data is in the interval $[0.9853, 6.1670]$. The results of this article strongly indicate that the estimated Kies style distributions are supported in intervals close to $(1, 9)$. This motivates us to transform the original data S to

$$S_{\text{new}} = \frac{S - \min(S)}{1.5\,(\max(S) - \min(S))}. \tag{6.3}$$

This way we can compare the current results with those presented in [52]. We devise now the interval $(0, 1)$ into $m = 20$ sub-intervals. The derived estimates are reported in the second part of Table 1 and the corresponding PDFs can be seen in Figure 3b. We can observe that the best fit produces the multimodal distribution – this is true for the first statistical sample too. This is not surprising since these distributions are very flexible and can fit different curves. On the other hand, they are prone to over-fitting and thus they should be used only when there is strong evidence that the considered statistical sample is indeed multimodal. The fit that produces the exponential mixture is remarkable given that the exponential distribution is driven by only one parameter. Of course, the gamma approximation is better since the gamma distribution generalizes the exponential one. Note that the geometric mixture produces a quite good fit too.

The average value of the empirical sample after the above-mentioned treatment is 0.1540. We can see in Table 1 that all models estimate relatively precisely the expectations. The reported confidence intervals are again stated rather at the left domain. The right endpoint is near 0.33 – only the beta mixture produces a larger value.

7. Asymptotic behavior of the estimator

Next, we discuss the consistency of the proposed estimator. We chose to use the exponential mixture defined in Section 5.2.1 because it depends only on two parameters – the power in the Kies distribution and the intensity of the exponential one. This way we can avoid the problems that arise from multidimensional optimization. We need the following lemma for the quantile function:

Lemma 7.1. *Suppose that the power coefficient β is a constant and λ is exponentially distributed with intensity θ. The quantile function of this mixture is*

$$q(x) = \frac{\theta^{\frac{1}{\beta}} x^{\frac{1}{\beta}}}{\theta^{\frac{1}{\beta}} x^{\frac{1}{\beta}} + (1-x)^{\frac{1}{\beta}}}. \tag{7.1}$$

Proof. We derive the cumulative distribution function through its complementary (5.19) as

$$F(t) = 1 - \overline{F}(t) = \frac{t^{\beta}}{\theta(1-t)^{\beta} + t^{\beta}} = \frac{1}{\theta\left(\frac{1}{t} - 1\right)^{\beta} + 1}. \tag{7.2}$$

It left to invert function (7.2). □

Our method for analyzing the estimator is based on the following steps:

1. We generate N uniformly distributed random numbers on the interval $(0, 1)$.
2. We apply to these numbers quantile function (7.1) with parameters β and θ – the result is a set with elements u_i, $i = 1, 2, ..., N$, distributed under the exponential Kies mixture.
3. We devise the set $(0, 1)$ into m bins and count the number N_i of u_i's falling in them. The PDF values at these nodes are obtained as $\frac{mN_i}{N}$ – see Section 6.

(4) We estimate the values of β and θ minimizing cost function (6.1). We use the MATLAB function `fminsearchbnd` developed by [11].
(5) We repeat this algorithm k times. Then we average the derived values as well as obtain the percent of deviations more than δ from the original β and θ. Also, we find the deviations by the proportion of δ from the original values.

We apply this approach with the following values: $\beta \in \{2, 3\}$, $\theta \in \{0.5, 1, 2, 4\}$, $k = 1\ 000$, $N \in \{1\ 000, 10\ 000, 100\ 000, 1\ 000\ 000\}$, $\epsilon = 10^{-5}$, and $\delta = 0.01$. The MATLAB function `fminsearchbnd` searches the minimum of the cost function in the set $\{(0, 10) \times (0, 10)\}$, the initial point is assumed to be the center. The results are reported in Table 2. It is divided into four parts – one for each value of N. The first two rows of each sector consist of the averaged values of β and θ. In the third and fourth rows are reported the percent of the estimations that deviate more than $\delta\,(= 0.01)$ from the original values. The last two lines are for the number of deviations larger than $\delta\beta$ and $\delta\theta$, respectively. The presented results lead to strong evidence that the estimator is consistent in the sense that the deviations from the original distribution tend to be zero. For the used MATLAB codes, see https://github.com/zhuszhus/Tsvetelin-Zaevski/tree/main.

In addition, in Figure 5 we present the estimations derived by the above-mentioned values of N and parameters $\beta = 3$ and $\theta = 4$. The estimated PDFs are in blue, the original ones are in red, and the values at the grid nodes are presented by green points. We can observe that both curves are indistinguishable for larger N's. The estimated parameters are reported in Table 3.

8. Conclusions and further work

A new mixture model based on the classical Kies distribution has been developed in this paper. The parameters that drive the original Kies distribution are assumed to be random variables. Several conditions that guarantee keeping the model behavior have been obtained. The probability properties of the resulting objects have been studied. Special attention is paid to the initial point of the probability density function. The necessary and sufficient conditions for it to be zero, finite, or infinite have been established. The so-called Hausdorff saturation is considered. It describes the position of the distributions with respect to the axes and thus it can be viewed as a measure of the rate of occurrence of a random event. Semi-closed form formulas are derived. Several particular examples have been studied. They are stated on bi- and multi-modal mixtures, other discrete mixtures (binomial and geometric), and mixtures based on continuous distributions – exponential, gamma, and beta. The proposed models are calibrated to two real-life data – one from the financial markets and another from the social sphere. The asymptotic behavior of the proposed estimator has been explored and its consistency has been checked.

Possible further investigations can be done in several directions. First, the relation between the Weibull and Kies distributions can be applied to defining many new models – the underlying fractional linear transform can be helpful in many practical tasks. Second, mixing different probability distributions is a powerful tool for combining their properties and thus increasing their applicability. Last but not least, the Hausdorff saturation can be used in many real-life fields – for example, to measure the failure rate (in engineering) or the financial risk (in risk management).

Acknowledgments

The authors would like to express sincere gratitude to the editor Prof. Nursel Koyuncu, the associated editor Prof. Haydar Demirhan, and the anonymous reviewers for the helpful and constructive comments which substantially improve the quality of this paper.

Author contributions. All the co-authors have contributed equally in all aspects of the preparation of this submission.

Conflict of interest statement. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding. The first author is financed by the European Union-NextGenerationEU, through the National Recovery and Resilience Plan of the Republic of Bulgaria, project No BG-RRP-2.004-0008. The second author is financed by the European Union-NextGenerationEU, through the National Recovery and Resilience Plan of the Republic of Bulgaria, project No BG-RRP-2.004-0001-C01.

Data availability. Open-access data is used in this research.

References

[1] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables,* US Government printing office, 1968

[2] T.A. Abushal, T.N. Sindhu, S.A. Lone, M.K.H. Hassan and A. Shafiq, *Mixture of Shanker Distributions: Estimation, Simulation and Application*, Axioms **12** (3), 231, 2023

[3] A.Z. Afify, A.M. Gemeay, N.M. Alfaer, G.M. Cordeiro and E.H. Hafez, *Power-Modified Kies-Exponential Distribution: Properties, Classical and Bayesian Inference with an Application to Engineering Data*, Entropy **24** (7), 883, 2022

[4] Z. Ahmad, E. Mahmoudi, R. Roozegar, M. Alizadeh and A.Z. Afify, *A new exponential-X family: modeling extreme value data in the finance sector*, Math. Probl. Eng. **2021** (1), 1-14, 2021

[5] A.A. Al-Babtain, M.K. Shakhatreh, M. Nassar and A.Z. Afify, *A new modified Kies family: Properties, estimation under complete and type-II censored samples, and engineering applications*, Mathematics **8** (8), 1345, 2020

[6] M.M. Al Sobhi, *The modified Kies–Fréchet distribution: properties, inference and application*, AIMS Math. **6**, 4691-4714, 2021

[7] A. Alsubie, *Properties and applications of the modified Kies-Lomax distribution with estimation methods*, J. Math. **2021** (1), 1-18, 2021

[8] M.H. Berger and D. Jeulin, *Statistical analysis of the failure stresses of ceramic fibres: dependence of the Weibull parameters on the gauge length, diameter variation and fluctuation of defect density*, J. Mater. Sci. **38** (13), 2913-2923, 2003

[9] F. Blasques, J. van Brummelen, P. Gorgi and S.J. Koopman, *Maximum likelihood estimation for non-stationary location models with mixture of Normal distributions*, J. Econom. **238** (1), 105575, 2024

[10] G. D'Amico, R. De Blasis and F. Petroni, *The mixture transition distribution approach to networks: evidence from stock markets*, Phys. A: Stat. Mech. Appl. **632**, 129335, 2023

[11] J. D'Errico, *fminsearchbnd.m; fminsearchcon.m*, https://www.mathworks.com/matlabcentral/fileexchange/8277-fminsearchbnd-fminsearchcon. MATLAB Central File Exchange, 2024

[12] F.S. dos Santos, K.K.F. do Nascimento, J. da Silva Jale, S.F.A. Xavier and T.A.E. Ferreira, *Brazilian wind energy generation potential using mixtures of Weibull distributions*, Renew. Sust. Energ. Rev. **189**, 113990, 2024

[13] W. Emam and Y. Tashkandy, *The Weibull claim model: Bivariate extension, Bayesian, and Maximum Likelihood estimations*, Math. Probl. Eng. **2022** (1), 8729529

[14] I.S. Gradshteyn and I.M. Ryzhik, *Table of ntegrals, Series, and Products*, Academic press, 2014

[15] M. Hashempour, *A weighted Topp-Leone G family of distributions: properties, applications for modelling reliability data and different method of estimation*, Hacet. J. Math. Stat. **51** (5), 1420-1441, 2022

[16] W. He, Z. Ahmad, A.Z. Afify and H. Goual, *The arcsine exponentiated-X family: validation and insurance application*, Complexity **2020** (1), 1-18, 2020

[17] Z. Hu, P. Li, D. Follmann and J. Qin, *A mixture distribution approach for assessing genetic impact from twin study*, Stat. Med. **41** (14), 2513-2522, 2022

[18] M. Khalid, M. Aslam and T.N. Sindhu, *Bayesian analysis of 3-components Kumaraswamy mixture model: Quadrature method vs. Importance sampling*, Alex. Eng. J. **59** (4), 2753-2763, 2020

[19] J.A. Kies, *The strength of glass performance*, Naval Research Lab Report **5093**, 1958

[20] S. Li, Y. Liu, Y. Sun and Y. Cai, *Deep learning-based channel estimation using Gaussian mixture distribution and expectation maximum algorithm*, Phys. Commun. **58**, 102018, 2023

[21] Y. Liu, D. Xie, D.A. Edwards and S. Yu, *Mixture copulas with discrete margins and their application to imbalanced data*, J. Korean Stat. Soc. **52** (4), 878-900, 2023

[22] S.A. Lone, S. Anwar, T.N. Sindhu and F. Jarad, *Some estimation methods for mixture of extreme value distributions with simulation and application in medicine*, Results Phys. **37**, 105496, 2022

[23] S.A. Lone, T.N. Sindhu, S. Anwar, M.K.H. Hassan, S.A. Alsahli and T.A. Abushal, *On construction and estimation of mixture of Log-Bilal distributions*, Axioms **12** (3), 309, 2023

[24] R. Maiboroda, V. Miroshnychenko and O. Sugakova, *Jackknife for nonlinear estimating equations*, Mod. Stoch.: Theory Appl. **9** (4), 377-399, 2022

[25] S. Matsushita, K. Hagiwara, T. Shiota, H. Shimada, K. Kuramoto and Y. Toyokura, *Lifetime data analysis of disease and aging by the Weibull probability distribution*, J. Clin. Epidemiol. **45** (10), 1165-1175, 1992

[26] M. Naderi and M.J. Nooghabi, *Clustering asymmetrical data with outliers: Parsimonious mixtures of contaminated mean-mixture of normal distributions*, J. Comput. Appl. Math. **437**, 115433, 2024

[27] D.E.Y. Sanku, M. Nassarn and D. Kumar, *Moments and estimation of reduced Kies distribution based on progressive type-II right censored order statistics*, Hacet. J. Math. Stat. **48** (1), 332-350, 2019

[28] C. Satheesh Kumar and S.H.S. Dharmaja, *On some properties of Kies distribution*, Metron **72** (1), 97-122, 2014

[29] C. Satheesh Kumar and S.H.S. Dharmaja, *The exponentiated reduced Kies distribution: properties and applications*, Commun. Stat. - Theory Methods **46** (17), 8778-8790, 2017

[30] C. Satheesh Kumar and S.H.S. Dharmaja, *On modified Kies distribution and its applications*, J. Stat. Res. **51** (1), 41-60, 2017

[31] J.V. Seguro and T.W. Lambert, *Modern estimation of the parameters of the Weibull wind speed distribution for wind energy analysis*, J. Wind Eng. Ind. Aerodyn. **85** (1), 75-84, 2000

[32] B. Sendov, *Hausdorff Approximations*, Springer Science & Business Media 50, 1990

[33] A. Shafiq, A.B. Çolak, C. Swarup, T.N. Sindhu and S.A. Lone, *Reliability analysis based on mixture of Lindley distributions with artificial neural network*, Adv. Theory Simul. **5** (8), 2200100, 2022

[34] A. Shafiq, S.A. Lone, T.N. Sindhu, Y. El Khatib, Q.M. Al-Mdallal and T. Muhammad, *A new modified Kies Fréchet distribution: applications of mortality rate of Covid-19*, Results Phys. **28**, 104638, 2021

[35] A. Shafiq, T.N. Sindhu, S.A. Lone, M.K.H. Hassan and K. Nonlaopon, *Mixture of Akash distributions: estimation, simulation and application*, Axioms **11** (10), 516, 2022

[36] A. Shafiq, A.B. Çolak, S.A. Lone, T.N. Sindhu and T. Muhammad, *Reliability modeling and analysis of mixture of Exponential distributions using artificial neural network*, Math. Meth. Appl. Sci. **47** (5), 3308-3328, 2024

[37] F.V.J. Silveira, F. Gomes-Silva, C.R. de Brito, J.S. Jale, F.R.S de Gusmão, S.F.A. Xavier-Junior and J.S. Rocha, *Modelling wind speed with a Univariate probability distribution depending on two baseline functions*, Hacet. J. Math. Stat. **52** (3), 808-827, 2023

[38] T.N. Sindhu, Z. Hussain and M. Aslam, *On the Bayesian analysis of censored mixture of two Topp-Leone distribution*, Sri Lankan J. Appl. Stat. **19** (1), 13-30, 2019

[39] T.N. Sindhu, Z. Hussain, N. Alotaibi and T. Muhammad, *Estimation method of mixture distribution and modeling of COVID-19 pandemic*, AIMS Math. **7** (6), 9926-9956, 2022

[40] A. Soulimani, M. Benjillali, H. Chergui and D.B. da Costa, *Multihop Weibull-fading communications: performance analysis framework and applications*, J. Frankl. Inst.-Eng. Appl. Math. **358** (15), 8012-8044, 2021

[41] M.T. Vasileva, *On Topp-Leone-G power series: saturation in the Hausdorff sense and applications*, Mathematics **11** (22), 4620, 2023

[42] M. Vasileva and N. Kyurkchiev, *Insuarance Mathematics*, Plovdiv University Press (in Bulgarian), 2023

[43] H. Wang, *Tolerance limits for mixture-of-normal distributions with application to COVID-19 data*, WIREs Comput. Stat. **15** (6), e1611, 2023

[44] Y. Wang, Z. Meng, Z. Zhang, M. Xia, L. Xia and W. Li, *A regularization algorithm of dynamic light scattering for estimating the particle size distribution of dual-substance mixture in water*, Particuology **89**, 246-257, 2024

[45] W. Weibull, *A statistical distribution function of wide applicability*, J. Appl. Mech. **18** (3), 293-297, 1951

[46] D.S. Wilks, *Rainfall intensity, the Weibull distribution, and estimation of daily surface runoff*, J. Appl. Meteorol. Climatol. **28** (1), 52-58, 1989

[47] A. Yan, J. Guo and D. Wang, *Robust stochastic configuration networks for industrial data modelling with Students-t mixture distribution*, Inf. Sci. **607**, 493-505, 2022

[48] J. Yazhou, W. Molin and J. Zhixin, *Probability distribution of machining center failures*, Reliab. Eng. Syst. Saf. **50** (1), 121-125, 1995

[49] Y.I. Yeleyko and O.A. Yarova, *Mixture of distributions based on the Markov chain*, Cybern. Syst. Anal. **58** (5), 754-757, 2022

[50] T.S. Zaevski and N. Kyurkchiev, *Some notes on the four-parameters Kies distribution*, C. R. Acad. Bulg. Sci. **75** (10), 1403-1409, 2022

[51] T. Zaevski and N. Kyurkchiev, *On some composite Kies families: distributional properties and saturation in Hausdorff sense*, Mod. Stoch.: Theory Appl. **10** (3), 287-312, 2023

[52] T. Zaevski and N. Kyurkchiev, *On min- and max-Kies families: distributional properties and saturation in Hausdorff sense*, Mod. Stoch.: Theory Appl. **11** (3), 265-288, 2024

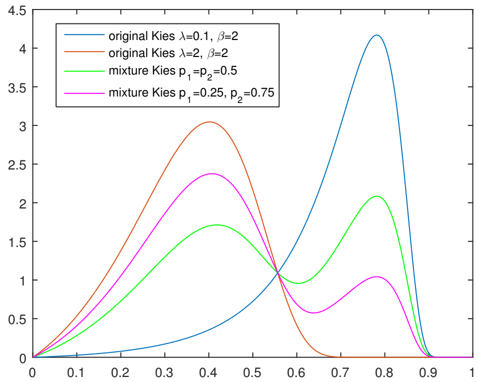
[53] T. Zaevski and N. Kyurkchiev, *On the Hausdorff saturation of some trigonometric-Kies families*, Palest. J. Math. **13** (2), 249-262, 2024

[54] Y. Zhang, Y. Dong and R. Feng, *Bayes-informed mixture distribution for the EVD estimation and dynamic reliability analysis*, Mech. Syst. Signal Proc. **197**, 110352, 2023

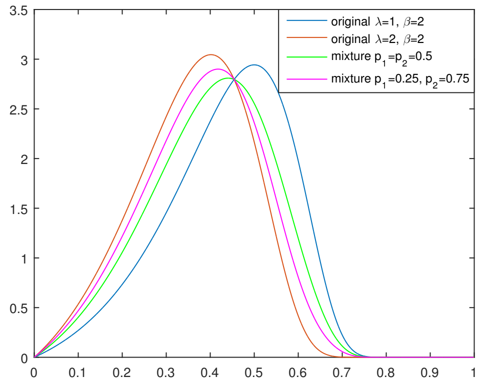
[55] Y. Zhenwu, Z. Ahmad, Z. Almaspoor and S.K. Khosa, *On the genesis of the Marshall-Olkin family of distributions via the T-X family approach: statistical modeling*, CMC-Comput. Mat. Contin. **67** (1), 753-760, 2021

Appendix A. PDFs of the linear transformed distributions

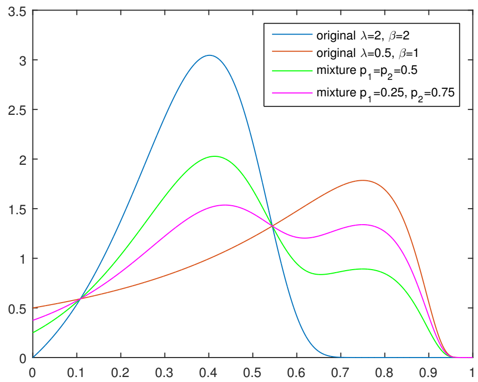
$$\begin{aligned}
f_{\mathrm{bin}}(t) &= \beta \frac{t^{\beta-1}}{(1-t)^{\beta+1}} \exp\left(-b\left(\frac{t}{1-t}\right)^{\beta}\right)\left(1-p+p\exp\left(-a\left(\frac{t}{1-t}\right)^{\beta}\right)\right)^{n-1} \times \\
&\quad \times \left[b-bp+p\,(na+b)\exp\left(-a\left(\frac{t}{1-t}\right)^{\beta}\right)\right] \\
f_{\mathrm{geo}}(t) &= p\beta t^{\beta-1} \frac{(b+a)\exp\left((a-b)\left(\frac{t}{1-t}\right)^{\beta}\right) - b\,(1-p)\exp\left(-b\left(\frac{t}{1-t}\right)^{\beta}\right)}{\left(\exp\left(a\left(\frac{t}{1-t}\right)^{\beta}\right)+p-1\right)^{2}(1-t)^{\beta+1}} \\
f_{\exp}(t) &= \frac{\theta\beta t^{\beta-1}(1-t)^{\beta-1}\exp\left(-b\left(\frac{t}{1-t}\right)^{\beta}\right)}{\left(\theta\,(1-t)^{\beta}+at^{\beta}\right)^{2}}\left[b\left(\theta+a\left(\frac{t}{1-t}\right)^{\beta}\right)+a\right] \qquad \text{(A.1)} \\
f_{\gamma}(t) &= \frac{\theta^{\alpha}\beta t^{\beta-1}(1-t)^{\alpha\beta-1}\exp\left(-b\left(\frac{t}{1-t}\right)^{\beta}\right)}{\left(\theta\,(1-t)^{\beta}+at^{\beta}\right)^{\alpha+1}}\left[b\left(\theta+a\left(\frac{t}{1-t}\right)^{\beta}\right)+a\alpha\right] \\
f_{\beta}(t) &= \beta\exp\left(-b\left(\frac{t}{1-t}\right)^{\beta}\right)\frac{t^{\beta-1}}{(1-t)^{\beta+1}} \\
&\quad \times \left[\begin{array}{l} b\,{}_1F_1\left(\alpha,\alpha+\theta,-a\left(\frac{t}{1-t}\right)^{\beta}\right) \\ +\frac{\alpha a}{\alpha+\theta}\,{}_1F_1\left(\alpha+1,\alpha+\theta+1,-a\left(\frac{t}{1-t}\right)^{\beta}\right) \end{array}\right]
\end{aligned}$$



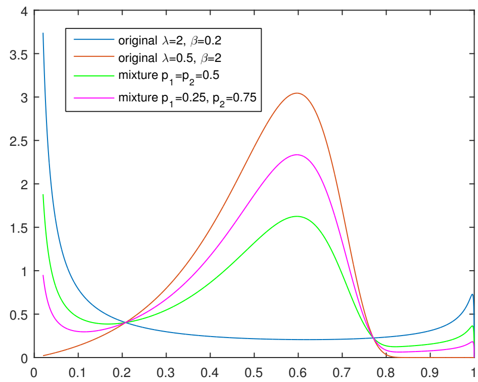
(a) $\lambda \in \{0.1, 2\}, \beta = 2$



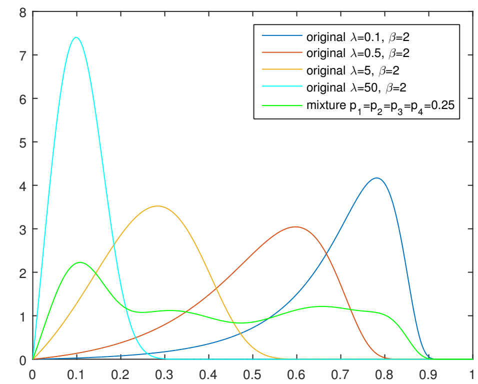
(b) $\lambda \in \{1, 2\}, \beta = 2$



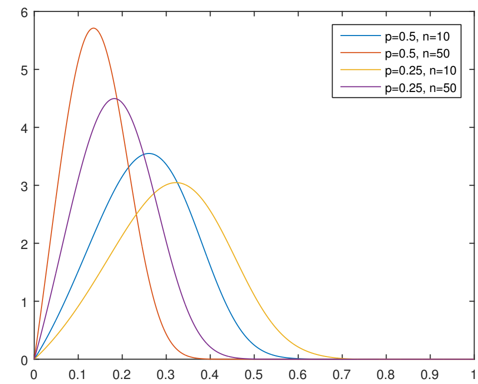
(c) $\lambda \in \{0.1, 2\}, \beta = 2$



(d) $\lambda \in \{0.1, 2\}, \beta = 2$

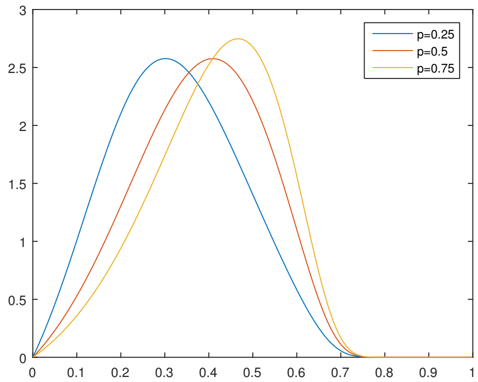


(e) $\lambda \in \{0.1, 2\}, \beta = 2$

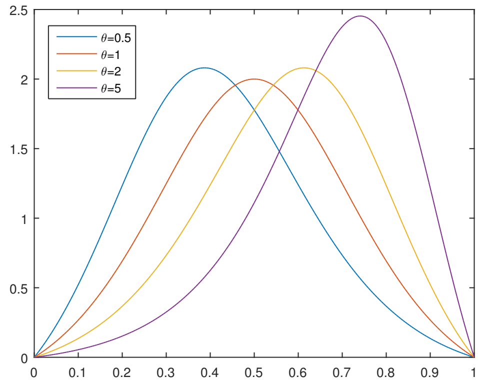


(f) Binomial mixture

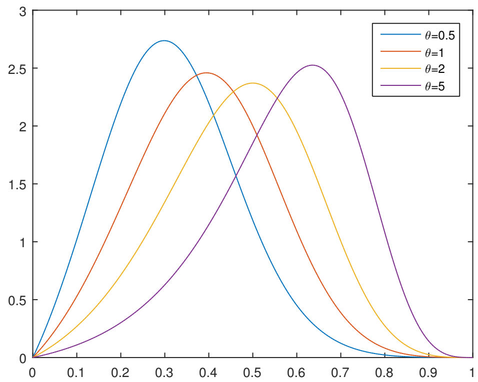
Figure 1. The figure contains the PDFs of different mixtures. The parameters are reported in the related legends and captions. The first four figures (a)-(d) are for the bi-modal mixtures, the PDFs of the original Kies distributions are presented too. A multi-modal mixture based on four original Kies distributions is depicted in Figure (e). Figure (d) presents four mixtures with different parameters for the binomial distributed λ; the parameter β is assumed to be deterministic and it is set to be $\beta = 2$.



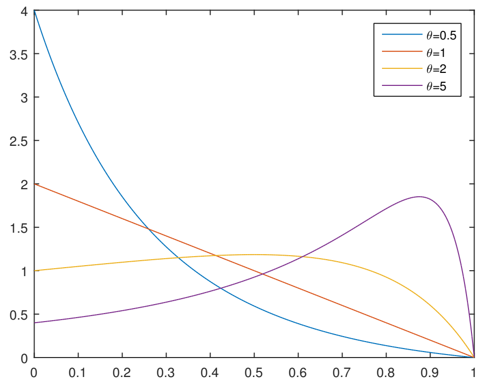
(a) Geometric mixture



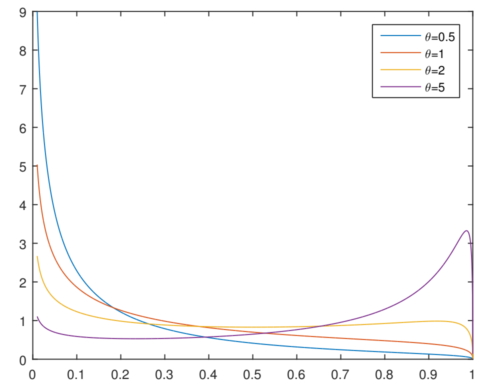
(b) Exponential mixture



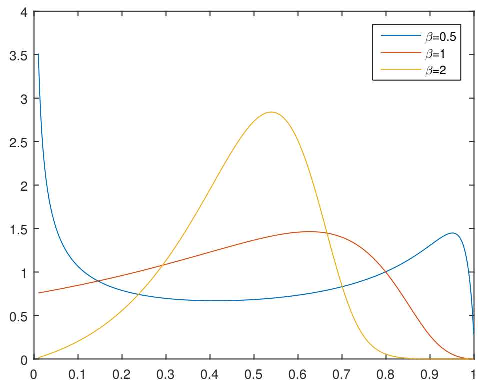
(c) Gamma mixture, $\beta = 2$



(d) Gamma mixture, $\beta = 1$

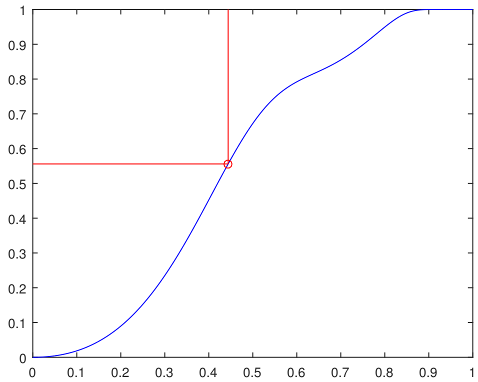


(e) Gamma mixture, $\beta = 0.7$

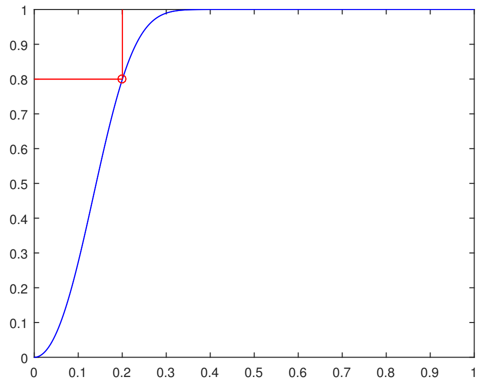


(f) Beta mixture

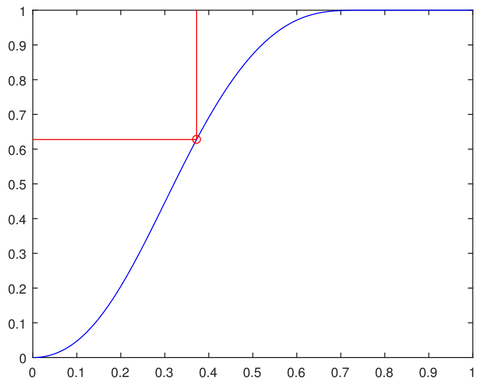
Figure 2. The figure contains the PDFs of four mixtures – one discrete and three continuous. The parameters are reported in the related legends and captions. For all of them, we assume that $\beta = 2$, except if a different value is given in the caption. Figure (a) provides three mixtures under the assumption that λ follows the geometric distribution. Figure (b) is for exponentially distributed λ – four intensities are considered. Figures (c)-(e) are for a gamma-distributed λ varying the power coefficient as $\beta \in \{0.7; 1; 2\}$. The parameter α of the gamma distribution is $\alpha = 2$ whereas θ is varied amongst $\theta \in \{0.5; 1; 2; 5\}$. Figure (f) is for the beta distributed random variable λ with parameters $(3, 1)$.



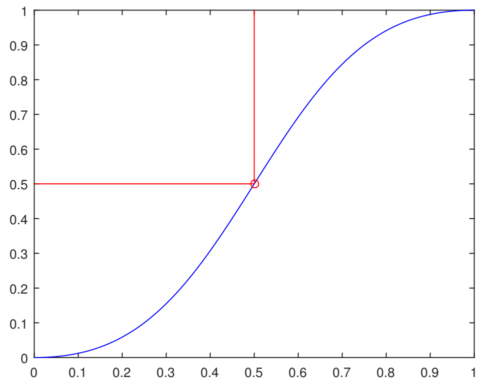
(a) Bimodal mixture



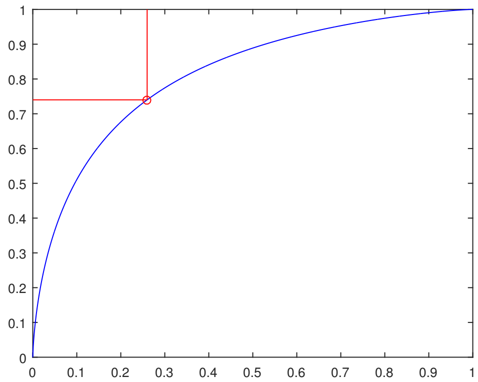
(b) Binomial mixture



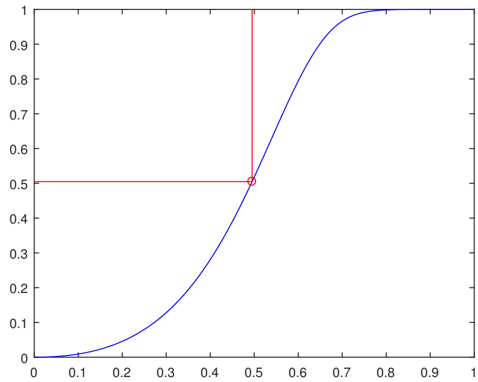
(c) Geometrical mixture



(d) Exponential mixture, $\beta = 1$

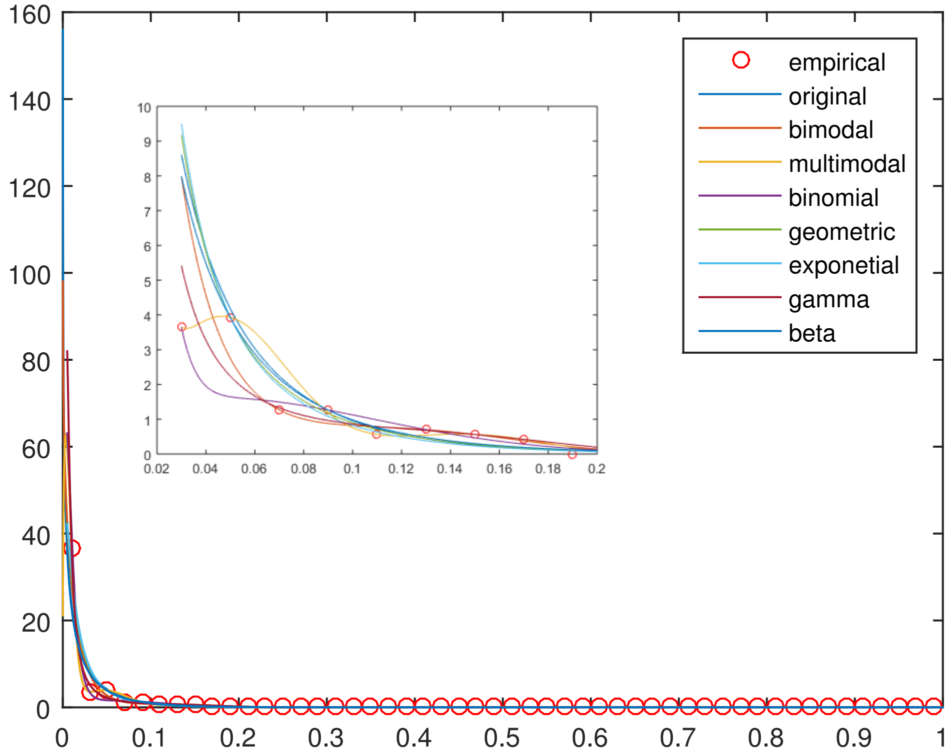


(e) Gamma mixture

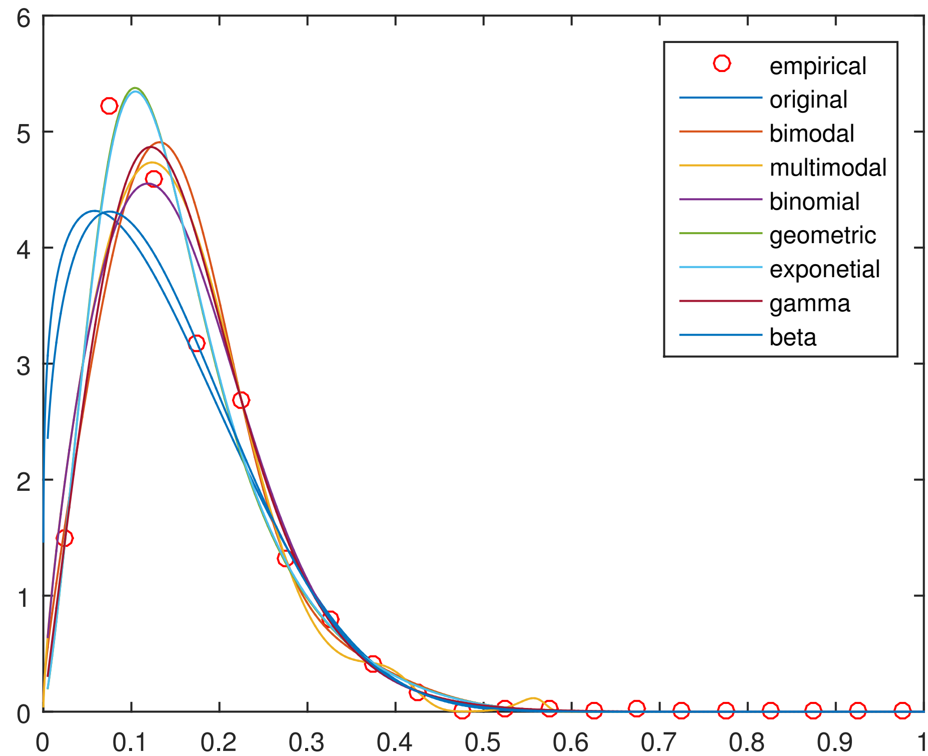


(f) Beta mixture

Figure 3. The figure presents the CDFs of six mixtures together with the Hausdorff saturations. They are depicted by red points. Note that the red lines and the zero point form squares. The power for the Kies family is assumed to be $\beta = 2$ except for the gamma mixture. Figure (a) is for a bimodal mixture with coefficients $p_1 = 0.25$ and $p_1 = 0.t5$; the parameters of the original Kies distributions are $\lambda = 0.1$ and $\lambda = 2$. In Figure (b) λ follows a binomial distribution with parameters $\{n = 50, p = 0.5\}$. Figure (c) is for a geometric distributed λ with parameter $p = 0.25$. Figure (d) is for an exponentially distributed λ with intensity $\theta = 1$. Figure (e) is for the gamma-mixture with parameters $\{\alpha = 2, \theta = 0.5\}$. The power coefficient is assumed to be $\beta = 0.7$. Figure (f) is for the beta-distributed λ with parameters $\{\alpha = 3, \theta = 1\}$.

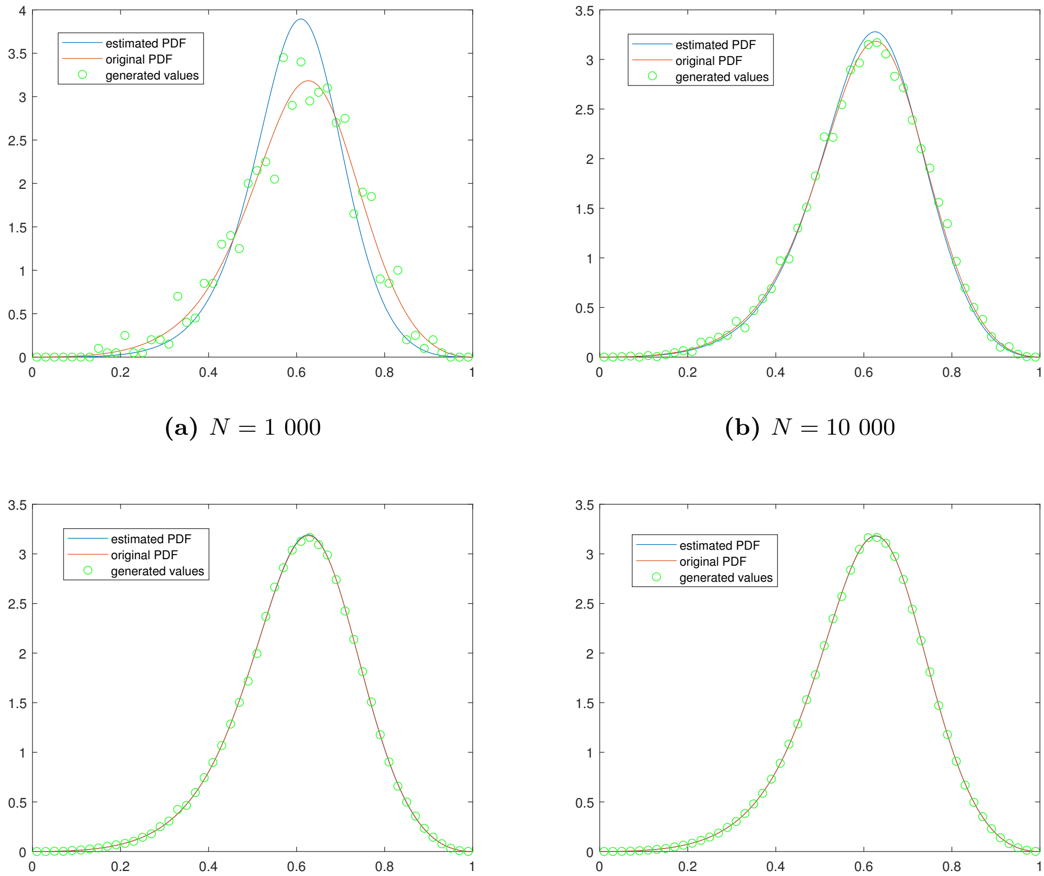


(a) S&P 500 index



(b) Unemployment Insurance Issues

Figure 4. The figure presents the PDFs of the real statistical samples as well as eight calibrated distributions – the original Kies and seven mixtures. Figure (a) is for the S&P 500 index, and Figure (b) is for the unemployment insurance issues.



(a) $N = 1\ 000$ **(b)** $N = 10\ 000$

(c) $N = 100\ 000$ **(d)** $N = 1\ 000\ 000$

Figure 5. Estimations of generated samples with sizes $N = 1\ 000$, $N = 10\ 000$, $N = 100\ 000$, and $N = 1\ 000\ 000$. The used distribution is an exponential mixture with power $\beta = 3$ and intensity $\theta = 4$.

Table 1. Numerical Estimations. The table presents the calibrated parameters for seven mixed distributions and the original Kies one. The resulting expectations and the confidence intervals are reported too. The first part is for the S&P 500 index, and the second one is for the unemployment insurance issues.

S&P 500	original	bimodal	multimodal	model	binomial	geometric	exponential	gamma	beta
	(A1)	(A2)	(A3)		(A4)	(A5)	(A6)	(A7)	(A8)
λ_1	15.7857	40.6027	735.5371	β	1.4975	1.1969	1.2479	3.2018	0.6653
λ_2	—	109.2334	920.2421	a	560.3198	20.8690	74.2504	732.5631	6.4268
λ_3	—	—	576.0060	b	33.0044	1.1514×10^{-8}	14.1627	185.3902	7.2682
β_1	0.7120	0.8796	1.3843	p	0.0017	0.1323	—	—	—
β_2	—	2.4893	3.9519	n	981	—	—	—	—
β_3	—	—	2.2741	θ	—	—	0.4193	1.6238×10^{-5}	4.0301×10^{-5}
p_1	1	0.9165	0.7175	α	—	—	—	0.2529	7.3923
p_2	—	0.0835	0.0494	—	—	—	—	— —	—
p_3	—	—	0.2331	—	—	—	—	— —	—
error	25.3497	22.8935	21.6932		23.0744	24.7371	25.1232	22.4339	24.5851
expectation	0.0241	0.0237	0.0241		0.0220	0.0246	0.0241	0.0234	0.0259
95% conf. int.	$(0, 0.0883)$	$(0, 0.1065)$	$(0, 0.1439)$		$(0, 0.1123)$	$(0, 0.0873)$	$(0, 0.0831)$	$(0, 0.0912)$	$(0, 0.0924)$
UII	(A1)	(A2)	(A3)		(A4)	(A5)	(A6)	(A7)	(A8)
λ_1	7.0550	27.4484	0.0245	β	1.6951	2.2257	2.2161	1.9603	1.2819
λ_2	—	5.0320	14.9210	a	3.8136	1.8680	4.7886	6.9343	3.9552
λ_3	—	—	16.9510	b	2.8160	0.9654	2.0540	0.3263	4.9723
β_1	1.1895	1.9935	16.3815	p	0.0334	0.0322	—	—	—
β_2	—	1.3558	1.6827	n	95	—	—	—	—
β_3	—	—	6.0798	θ	—	—	0.0869	0.6134	0.0003
p_1	1	0.6464	0.0048	α	—	—	—	2.1985	0.0011
p_2	—	0.3536	0.9584	—	—	—	—		—
p_3	—	—	0.0369	—	—	—	—		—
error	7.1362	6.1300	2.8820		5.6375	5.8140	5.7291	5.4682	6.6078
expectation	0.1418	0.1606	0.1550		0.1586	0.1546	0.1549	0.1611	0.1447
95% conf. int.	$(3.9063 \times 10^{-6}, 0.3274)$	$(0.0059, 0.3373)$	$(0.0044, 0.3228)$		$(0.0061, 0.3307)$	$(0.0126, 0.3388)$	$(0.0126, 0.3387)$	$(0.0115, 0.3334)$	$(7.8125 \times 10^{-6}, 0.3815)$

Table 2. Consistency of the Estimator. The table reports the average of the calibrated parameters of the exponential Kies mixture as well as the number of large deviations. The data consists of generated random numbers distributed through this law. The power-intensity pair $\{\beta, \theta\}$ is amongst eight alternatives. The four sections are for data size $N \in \{1\ 000, 10\ 000, 100\ 000, 1\ 000\ 000\}$.

	$\{\beta,\theta\}$	$\{\beta,\theta\}$	$\{\beta,\theta\}$	$\{\beta,\theta\}$	$\{\beta,\theta\}$	$\{\beta,\theta\}$	$\{\beta,\theta\}$	$\{\beta,\theta\}$
$N = 1\ 000$	$\{2, 0.5\}$	$\{2, 1\}$	$\{2, 2\}$	$\{2, 4\}$	$\{3, 0.5\}$	$\{3, 1\}$	$\{3, 2\}$	$\{3, 4\}$
$\overline{\beta}$	2.0492	2.0505	2.0517	2.0824	3.5725	3.6111	3.5636	3.5380
$\overline{\lambda}$	0.4874	1.0048	2.0578	4.3601	0.4392	1.0048	2.3077	5.1257
dev. β, 0.01	94.9	95.2	94.8	94.9	99.8	100	99.8	100
dev. of β from 1%	87.5	90.1	86.9	89.4	99.8	99.8	99.4	99.9
dev. λ, 0.01	82.9	84.8	95.3	98.6	92.6	86	97.8	99.8
dev. of λ from 1%	91.9000	84.8	92.1	93.1	96.2	86	95.8	98.9
$N = 10\ 000$	$\{2, 0.5\}$	$\{2, 1\}$	$\{2, 2\}$	$\{2, 4\}$	$\{3, 0.5\}$	$\{3, 1\}$	$\{3, 2\}$	$\{3, 4\}$
$\overline{\beta}$	2.0037	2.0016	2.0031	2.0043	3.0525	3.0486	3.0513	3.0569
$\overline{\lambda}$	0.4993	1.0003	2.0039	4.0190	0.4936	0.9994	2.0291	4.1297
dev. β, 0.01	65.4	65.8	64	62	89.4	86.8	91	90.4
dev. of β from 1%	40.6	40.1	41	35	73	65.3	73.6	71.1
dev. λ, 0.01	44.7000	67.3	83.1	92.8	53.3	71.1	87.7	95.5
dev. of λ from 1%	69.8	67.3	68.6	73.4	74.7	71.1	76	83.6
$N = 100\ 000$	$\{2, 0.5\}$	$\{2, 1\}$	$\{2, 2\}$	$\{2, 4\}$	$\{3, 0.5\}$	$\{3, 1\}$	$\{3, 2\}$	$\{3, 4\}$
$\overline{\beta}$	1.9993	1.9994	1.9994	1.9994	2.9996	2.9960	2.9984	2.9994
$\overline{\lambda}$	0.5000	1.0003	2.0000	3.9989	0.4996	1.0001	2.0015	4.0061
dev. β, 0.01	21.2	23.8	22.7	21	52	55.6	50	59.2
dev. of β from 1%	1.3	2	1.9	1.9	8.9	11.5	7.8	11
dev. λ, 0.01	1.7	19.7	54.5	76.6	4.6	28.2	59.6	83.4
dev. of λ from 1%	22.7	19.7	21.9	28.9	31.2	28.2	27.2	43.2
$N = 1\ 000\ 000$	$\{2, 0.5\}$	$\{2, 1\}$	$\{2, 2\}$	$\{2, 4\}$	$\{3, 0.5\}$	$\{3, 1\}$	$\{3, 2\}$	$\{3, 4\}$
$\overline{\beta}$	1.9989	1.999	1.9991	1.9992	2.9930	2.9935	2.9931	2.9931
$\overline{\lambda}$	0.5002	1.0001	1.9992	3.9977	0.5005	1.0000	1.9982	3.9924
dev. β, 0.01	0	0.2	0	0.1	33.0044	26.5	27.6	28.6
dev. of β from 1%	0	0	0	0	0	0	0	0
dev. λ, 0.01	0	0	6.5	40.2	0	0.1	10	56.2
dev. of λ from 1%	0	0	0	0.1	0	0.1	0.2	2.2

Table 3. The table reports the calibrated parameters of an exponential Kies mixture. The sample consists of generated through this law random numbers – with power parameter $\beta = 3$ and intensity $\theta = 4$. The data size is amongst $N \in \{1\ 000, 10\ 000, 100\ 000, 1\ 000\ 000\}$.

	$N = 1\ 000$	$N = 10\ 000$	$N = 100\ 000$	$N = 1\ 000\ 000$
β	3.7200	3.0925	3.0088	2.9956
λ	4.6673	4.1557	4.0064	3.9978