RT&A, No 4 (76)
Volume 18, December 2023

EXPONENTIAL-PARETO MIXTURE DISTRIBUTION

Irina Peshkova

•

Petrozavodsk State university,
Institute of Applied Mathematical Research of the Karelian Research Centre of RAS
iaminova@petrsu.ru

Abstract

In this paper we introduce the Exponential-Pareto mixture distribution. This distribution is associated as mixture of light and heavy-tailed data which arise in a wide class applications including risk analysis. Characteristic function, failure rate function, mean excess, conditional excess distribution are derived. It is proved that the limiting distribution of maxima among n values of rv's with Exponential-Pareto distribution has Frechet-type form. The maximal likelihood estimation of parameters is discussed. The upper bound of uniform distance between Exponential-Pareto mixture and Pareto distributions is derived.

Keywords: finite mixture, Exponential-Pareto distribution

1. Introduction

Finite mixture models are often used in the modelling time to failure of the systems in the competing risk situations. The exponential, Weibull, lognormal and Pareto distributions occupy a central role because of their demonstrated usefulness in the analysis of lifetime data and in problems related to the modelling of ageing or failure processes. Finite mixtures are also useful in medical and biology research, in artificial neural networks and robustness studies, income analysis [1].

In lifetime data analysis, the population of lifetimes usually can be decomposed into subpopulations. Moreover, The data may be actually generated by quite different distributions, in particular, distributions with the so-called *light or heavy tails*. For example, an insurance portfolio may include both many small (light) claims and also a few large (heavy) claims. Hence the claims distribution can be modeled as a finite mixture distributions with different tail behavior. [2] In this paper we propose to consider a mixture of exponential distribution and the Pareto distribution to model such situations since the Pareto distribution relates to the class of long-tailed distributions, while the exponential distribution is light-tailed.

The exponential distribution appeared suitable for modelling the lifetimes of various types of manufactured items. The Pareto model is applied in many fields, for example, in income analysis, in signal processing for simulation of X-band maritime surveillance radar clutter [3].

The mixtures of light and heavy-tailed distribution can be useful in the simulation of IoT traffic, because, as it is known, the smart-home and smart-city environments can generate both short and large-sized packets. These environments involve several sensors dedicated to specific tasks, such as monitoring systems or collecting cyber-physical values (temperature, humidity, etc.). Smart cameras generate continuous data flows with large-sized packets, while smart plugs generate small-sized packets at a slow pace [4].

Various issues of the Exponential-Pareto mixture were considered by authors in the papers [5, 6, 7] in relation to queueing systems.

The main contribution of the paper is to introduce the Exponential-Pareto mixture distribution. It is the two-component mixture of exponential and Pareto distributions, which is used to approximate mixed data with light and heavy tails.

The paper is organized as follows. The moments and characterization function are derived in Section 2. The subexponentiality of the distribution is proved in Section 2. In Section 4 we discuss the properties of the conditional excess distribution. In Section 5 the extreme behavior of the Exponential-Pareto mixture distribution is discussed. In Section 6 We apply the known upper bound for Kullback-Leibler divergence between exponential and Pareto distributions [3, 8] to derive the convergence of uniform distance between Exponential-Pareto mixture and exponential distributions to zero as shape parameter goes to infinity. The estimation of parameters by log-likelihood maximization is discussed in Section 7..

2. Definition and moments of Exponential-Pareto mixture distribution

Definition 1. We say that r.v. Z has an *Exponential-Pareto mixture distribution* if its distribution function (df) has the following form:

$$F_Z(x) = 1 - pe^{-\lambda x} - (1-p)\left(\frac{x_0}{x_0 + x}\right)^{\alpha},\ \lambda > 0, \alpha > 0, x_0 > 0, x \geq 0, \tag{1}$$

where $0 < p < 1$ is *mixing proportion*.

Equation (1) shows that r.v. Z coincides with exponential distribution with the probability p, and with Pareto distribution with the probability $1-p$.The above can be reformulated as follows. Suppose that the random variables X, Y with distribution functions F_X, F_Y, respectively, are independent, and let I be indicator function independent of X, Y, taking value 1 with probability p (value 0 with probability $1-p$). Then it is said that the variable

$$Z = I\,X + (1-I)\,Y \tag{2}$$

has *two-component mixture distribution*. If X has exponential distribution $Exp(\lambda)$ and Y has Pareto distribution $Pareto(\alpha, x_0)$, then df of Z is given by (1).

The density function of Exponential-Pareto mixture distribution has form

$$f_Z(x) = p\lambda e^{-\lambda x} + (1-p)\frac{\alpha x_o^{\alpha}}{(x_0 + x)^{\alpha+1}}.$$

and has the following limits

$$\begin{aligned}\lim_{x\to\infty} f_Z(x) &= 0;\\ \lim_{x\to 0} f_Z(x) &= p\lambda + \frac{(1-p)\alpha}{x_0}.\end{aligned}$$

Figure 1 depicts cdf of Exponential-Pareto mixture distribution with parameters $x_0 = 0.5, \lambda = 2, p = 0.5$ and varied $\alpha = 0.5; 1.5; 2.5$.

Lemma 1. Let rv Z follow the Exponential-Pareto mixture distribution, then its characteristic function is given by

$$\phi_Z(t) = \frac{\lambda\alpha}{\lambda - ipt} e^{-i(1-p)tx_0} \sum_{k=0}^{\infty} \frac{(i(1-p)tx_0)^k}{k!(\alpha - k)} \tag{3}$$

Proof. Substitute df (1) into the formula for characteristic function $\phi_X(t)$ of rv Z

$$\phi_Z(t) = e^{it(pX + (1-p)Y)} = e^{pitX} e^{(1-p)itY} = \phi_X(pt)\phi_Y((1-p)t).$$

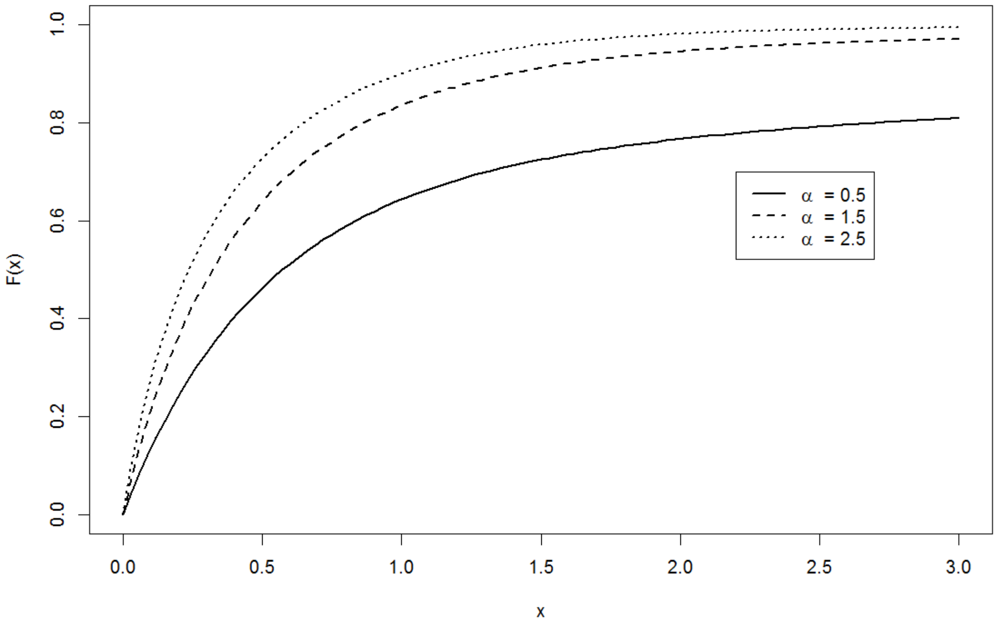

Figure 1: *Cumulative distribution function (cdf) of Exponential-Pareto distribution.*

where the characteristic function of first component X is given by

$$\phi_X(t) = \int_0^\infty e^{itx} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - it}.$$

The characteristic function of Pareto distribution can be calculated using the expansion of the exponential function in a Taylor series in the vicinity of zero

$$\begin{aligned}
\phi_Y(t) &= \int_0^\infty e^{itx} \alpha x_0^\alpha (x_0 + x)^{-\alpha-1} dx \\
&= e^{-itx_0} \int_{x_0}^\infty e^{itx} x_0^\alpha \alpha x^{-\alpha-1} dx \\
&= e^{-itx_0} x_0^\alpha \alpha \int_{x_0}^\infty \sum_{k=0}^\infty \frac{(itx)^k}{k!} x^{-\alpha-1} dx \\
&= e^{-itx_0} x_0^\alpha \alpha \sum_{k=0}^\infty \frac{(it)^k}{k!} \int_{x_0}^\infty x^{k-\alpha-1} dx \\
&= e^{-itx_0} \alpha \sum_{k=0}^\infty \frac{(itx_0)^k}{k!(\alpha - k)}.
\end{aligned}$$

Hence, the characteristic function of Exponential-Pareto mixture distribution takes form (3). ■

Lemma 2. Let rv Z follow the Exponential-Pareto mixture distribution, then its moments are given by

$$\mathbf{E}Z^k = \Gamma(1+k)\left(\frac{p}{\lambda}\right)^k + \frac{\Gamma(1+k)x_0^k(1-p)^k}{\Gamma(\alpha)} \sum_{j=0}^{k-1} \left(\frac{p}{x_0\lambda(1-p)}\right)^j \Gamma(\alpha - k + j),\ \alpha > k,$$

where $\Gamma(z) = \int\limits_0^\infty t^{z-1}e^{-t}dt$ is Gamma function.

The proof obviously follows from the well-known relation between moments and characterization function

$$i^{-k}\frac{d^k\phi_Z(t)}{dt^k}|_{t=0} = \mathbf{E}Z^k,$$

In particular, the mean is

$$\mathbf{E}Z = \frac{p}{\lambda} + \frac{(1-p)x_0}{\alpha-1},\ \alpha > 1,$$

the second moment is

$$\mathbf{E}Z^2 = \frac{2p^2}{\lambda^2} + \frac{2p(1-p)x_0}{\lambda(\alpha-1)} + \frac{2(1-p)^2x_0^2}{(\alpha-1)(\alpha-2)},\ \alpha > 2,$$

and variance is

$$\mathbf{Var}Z = \frac{p^2}{\lambda^2} + \frac{(1-p)^2x_0^2\alpha}{(\alpha-1)^2(\alpha-2)},\ \alpha > 2.$$

Note that moments of order k exist only for $\alpha > k$.

3. Failure rate function and subexponentiality

Consider the equilibrium df F_e of rv Z with df (1)

$$F_e(x) = \frac{1}{\mathbf{E}Z}\int\limits_0^x \overline{F}_Z(t)dt = 1 - \frac{1}{\mathbf{E}Z}\left(\frac{pe^{-\lambda x}}{\lambda} + \frac{(1-p)x_0^\alpha}{(\alpha-1)(x_0+x)^{\alpha-1}}\right), \tag{4}$$

where $\overline{F}(x) = 1 - F(x)$ is the tail of df F. We emphasize that expression (4) exists if parameter $\alpha > 1$.

We calculate the failure rate function of the Exponential-Pareto mixture distribution

$$r(x) := \frac{f_Z(x)}{\overline{F}_Z(x)} = \frac{p\,\lambda a(x) + (1-p)\,\alpha/(x_0+x)}{p\,a(x) + (1-p)}, \tag{5}$$

where

$$a(x) = e^{-\lambda x}\left(1 + \frac{x}{x_0}\right)^\alpha. \tag{6}$$

is an auxiliary function.

Note that

$$r(x) \longrightarrow 0 \quad \text{as} \quad x \to \infty,$$

which is typical for long-tailed distributions. Recall that df F_Z is *long-tailed*, if for each fixed $x > 0$ the following relation holds

$$\lim_{u\to\infty}\mathbf{P}(X > u + x|X > u) = 1.$$

This asymptotic property means, that for each fixed $x > 0$, the random variable Z exceeds the threshold $x + u$ with probability approaching 1 as u increasing. Long-tailed distributions have asymptotically decaying to zero failure rate functions (not necessary monotone) [9]. Below we prove that Exponential-Pareto distribution belongs to a subclass of long-tailed distribution – so-called *subexponential* distributions.

Note, that the failure rate function has the same limit as density function as $x \to 0$,

$$r_Z(x) \longrightarrow p\lambda + (1-p)\frac{\alpha}{x_0} \quad \text{as} \quad x \to 0.$$

Provided that the following relation between the parameters of the mixture is satisfied $\lambda \geq \alpha/x_0$, failure rate function is bounded from above, $r_Z(x) \leq \lambda$.

To verify the monotonicity of the failure rate function we calculate its derivative

$$\frac{dr_Z(x)}{dx} = -(1-p)\frac{pa(x)(r_X(x)-r_Y(x))^2 + r_Y^2(x)/\alpha(pa(x)+(1-p))}{(pa(x)+(1-p))^2} < 0,$$

Since $\frac{dr_Z(x)}{dx}$ is negative for all x, then $r_Z(x)$ decreases monotonically and, therefore, the Exponential-Pareto distribution belongs to the class of distributions with decreasing failure rate functions.

The following lemma states that the distribution (1) and the equilibrium distribution (4) are both belong to the class of so-called subexponential distributions. Recall that df F_Z is called *subexponential* [9] if

$$\lim_{x\to\infty} \frac{\overline{F_Z^{*n}}(x)}{n\overline{F}_Z(x)} = 1 \ \text{ for all } n \geq 2,$$

where $\overline{F_Z^{*n}}(x)$ is the tail of n-convolution of the distribution $F_Z(x)$ with itself.

To verify that the Exponential-Pareto mixture distribution F_Z and the corresponding equilibrium distribution F_e (with parameter $\alpha > 1$) both belong to the class of the subexponential distributions, it is enough [9] to verify that F_Z belongs to a special subclass $\mathcal{S}^*$ of the subexponential distributions.

Lemma 3. The Exponential-Pareto mixture distribution with df defined by expression (1) (with parameter $\alpha > 1$) belongs to a special subclass $\mathcal{S}^*$ of the subexponential distributions.

Proof. One of the following criteria for df to belong $\mathcal{S}^*$ can be applied [9].

1. If
$$\limsup_{x\to\infty} xr(x) < \infty, \tag{7}$$
then $F_Z \in \mathcal{S}^*$.

2. Suppose that
$$\lim_{x\to\infty} r(x) = 0. \tag{8}$$
Then
$$F_Z \in \mathcal{S}^* \iff \lim_{x\to\infty} \int_0^x e^{yr(x)}\overline{F}_Z(y)dy = \mathbf{E}Z. \tag{9}$$

It is easy to check that for Exponential-Pareto mixture distribution

$$r(x) \longrightarrow 0 \quad \text{as} \quad x \to \infty \ \text{ and } \ xr(x) \longrightarrow \alpha \quad \text{as} \quad x \to \infty.$$

Moreover,

$$\begin{aligned}\int_0^x e^{yr(x)}\overline{B}(y)dy \quad &= \quad -\frac{pe^{-y(\lambda - r(x))}}{\lambda - r(x)}\Big|_0^x + (1-p)\frac{(x_0 r(x))^\alpha}{r(x)}e^{-x_0 r(x)} \cdot \\ &\cdot \quad [\gamma(-\alpha+1, x_0 r(x) + xr(x)) - \gamma(-\alpha+1, x_0 r(x))] \longrightarrow \\ &\longrightarrow \quad \frac{p}{\lambda} + \frac{(1-p)x_0}{\alpha-1} = \mathbf{E}Z \quad \text{as} \quad x \to \infty\end{aligned}$$

since

$$\gamma(-\alpha+1, x_0 r(x) + xr(x)) \longrightarrow \gamma(-\alpha, \alpha) \quad \text{as} \quad x \to \infty$$

and

$$\gamma(-\alpha+1, x_0 r(x)) \sim \frac{(xr(x))^{-\alpha+1}}{-\alpha+1} \quad \text{as} \quad x \to \infty,$$

where $\gamma(\alpha, x) = \int_0^x y^{\alpha-1}e^{-y}dy$ is the lower incomplete gamma function.

Hence conditions (7)-(9) are satisfied and Exponential-Pareto distribution belongs to subclass $\mathcal{S}^*$. ■

4. Conditional Excess Distribution

In this section we consider conditional excess distribution of Exponential-Pareto mixture distribution.

On the event $\{Z > u\}$, define the excess $Z_u := Z - u$, which has the *conditional excess distribution over the threshold u* [10]

$$P(Z_u \le x) = P(Z - u \le x | Z > u),\ u \ge 0,\ \ x \ge 0. \tag{10}$$

Let $X_u,\ Y_u$ be conditional excesses of rv's X and Y with conditional excess distributions defined by (10), relatively. In the case of Exponential-Pareto mixture distribution expression for the tail of conditional excess distribution becomes

$$\overline{F}_{Z_u}(x) = \frac{\overline{F}_Z(x+u)}{\overline{F}_Z(u)} = \frac{pe^{-\lambda(x+u)} + (1-p)\left(\frac{x_0}{x_0+x+u}\right)^{\alpha}}{pe^{-\lambda u} + (1-p)\left(\frac{x_0}{x_0+u}\right)^{\alpha}}.$$

Note that rv Z_u is not a mixture of rv's X_u and Y_u, namely $Z_u \neq IX_u + (1-I)Y_u$.

The expected value of conditional excess is given by *the mean excess function*, defined as [10]

$$\begin{aligned} e_Z(u) = \mathbf{E}Z_u &= \int_0^\infty \frac{\overline{F}_Z(x+u)}{\overline{F}_Z(u)}dx = \frac{\int_u^\infty \overline{F}_Z(x)dx}{\overline{F}_Z(u)} \\ &= \frac{\int_u^\infty (pe^{-\lambda x} + (1-p)x_0^\alpha (x_0+x)^{-\alpha})dx}{pe^{-\lambda x} + (1-p)x_0^\alpha (x_0+x)^{-\alpha}} \\ &= \frac{pe^{-\lambda u}(\alpha-1)(x_0+u)^\alpha + \lambda(1-p)x_0^\alpha(x_0+u)}{\lambda p(\alpha-1)(x_0+u)^\alpha e^{-\lambda u} + \lambda(1-p)(\alpha-1)x_0^\alpha}.\ \ \alpha > 1. \end{aligned}$$

Note that the mean excess function is an increasing function, $e_Z(u) \to \infty$ as $u \to \infty$. The Failure rate function of Z_u is given by

$$r_{Z_u}(x) = \frac{p\, r_{X_u}(x) a(x+u) + (1-p)\, r_{Y_u}(x+u)}{p\, a(x+u) + (1-p)} = r_Z(x+u), \tag{11}$$

where $a(x)$ is defined by expression (6).

Now we derive the condition on the parameters of the Exponential-Pareto mixture distribution that guarantees the mixture to be bounded by its components in terms of the failure rate ordering. To prove it we need the following lemma.

Lemma 4. Let $r_X,\ r_Y$ be the failure rates of F_X, F_Y, respectively. Then if

$$\sup_{x\ge 0} r_X(x) < \infty \ \text{ and } \ \inf_{x \ge 0} r_Y(x) > 0 \tag{12}$$

and

$$\sup_{x\ge 0} r_X(x) \le \inf_{x\ge 0} r_Y(x), \tag{13}$$

then

$$X \underset{r}{\ge} Y,$$

where notation $X \underset{r}{\ge} Y$ means that X is *more than Y in failure rate*, i.e. $r_X(x) \le r_Y(x)$ for all x.

Theorem 1. Let rv Z have Exponential-Pareto mixture distribution with df defined by (1). If the following inequality holds

$$\frac{\alpha}{x_0} \leq \lambda, \tag{14}$$

then

$$X \underset{r}{\leq} Z \underset{r}{\leq} Y, \tag{15}$$

$$X_u \underset{r}{\leq} Z_u \underset{r}{\leq} Y_u, \tag{16}$$

$$e_X(u) \leq e_Z(u) \leq e_Y(u) \tag{17}$$

Proof. To prove the theorem, it suffices to find conditions under which the relations

$$r_Y(x) \leq r_Z(x) \leq r_X(x), \tag{18}$$

are satisfied. From lemma 4 for this it is sufficient that the following relation be satisfied:

$$r_Y(x) \leq \sup_{x \geq 0} r_Y(x) = r_Y(0) = \frac{\alpha}{x_0} \leq r_X(x) = \lambda. \tag{19}$$

Inequality (19) can be rewrite in form (14) whence it follows (15).

From statement 2.1 in [11] it follows that if rv's are ordered in failure rate, then their conditional excesses are ordered in failure rate too, that proves the statement (16).

Now we calculate the mean excess functions for rv's X and Y. Easy to check, that for exponential distribution mean excess function is equal to mathematical expectation, $e_X(u) = \mathbf{E}X = 1/\lambda$. For Pareto distribution we have

$$e_Y(u) = \frac{x_0 + u}{\alpha - 1}.$$

Obviously, if the condition of the theorem (14) is fulfilled, then inequality

$$\lambda \geq \frac{\alpha}{x_0} \geq \frac{\alpha - 1}{x_0 + u}$$

holds and

$$\frac{1}{\lambda} \leq \frac{pe^{-\lambda u}(\alpha-1)(x_0+u)^\alpha + \lambda(1-p)x_0^\alpha(x_0+u)}{\lambda p(\alpha-1)(x_0+u)^\alpha e^{-\lambda u} + \lambda(1-p)(\alpha-1)x_0^\alpha} \leq \frac{x_0+u}{\alpha-1},$$

that proves the relation (17) of the theorem. ■

Figure 2 demonstrates the failure rate functions of rv's X, Y, Z, where $\lambda = 1, \alpha = 0.5, x_0 = 0.6$. It can be seen from the graph that, for the given parameters, the failure rate functions are ordered in accordance with the relation (18), hence it follows from Theorem 1 that rv's are ordered in accordance with (17).

Figure 3 shows mean excess functions of rv's X, Y, Z with $\lambda = 5, \alpha = 2.1, x_0 = 0.5, p = 0.5$. It can be seen from the graph that, for the given parameters, these functions are ordered as in (17).

5. Extreme behavior of the Exponential-Pareto distribution

Let $\{X_n,\ n \geq 1\}$ be a family of the independent and identically distributed (iid) rv's with a distribution function F. Then the distribution of $M_n = \max(X_1, \ldots, X_n)$ satisfies $\mathbf{P}(M_n \leq x) = F^n(x)$.

Suppose there exists a sequence of real constants $b_n,\ a_n > 0,\ n \geq 1$ such that

$$\lim_{n\to\infty} \mathbf{P}((M_n - b_n)/a_n \leq x) = G(x), \quad n \to \infty, \tag{20}$$

for every continuity point x of G, and G a nondegenerate distribution function. Then $G(x)$ is one of the three types of *extreme value distributions*: Gumbel, Frechet or reversed Weibull [12].

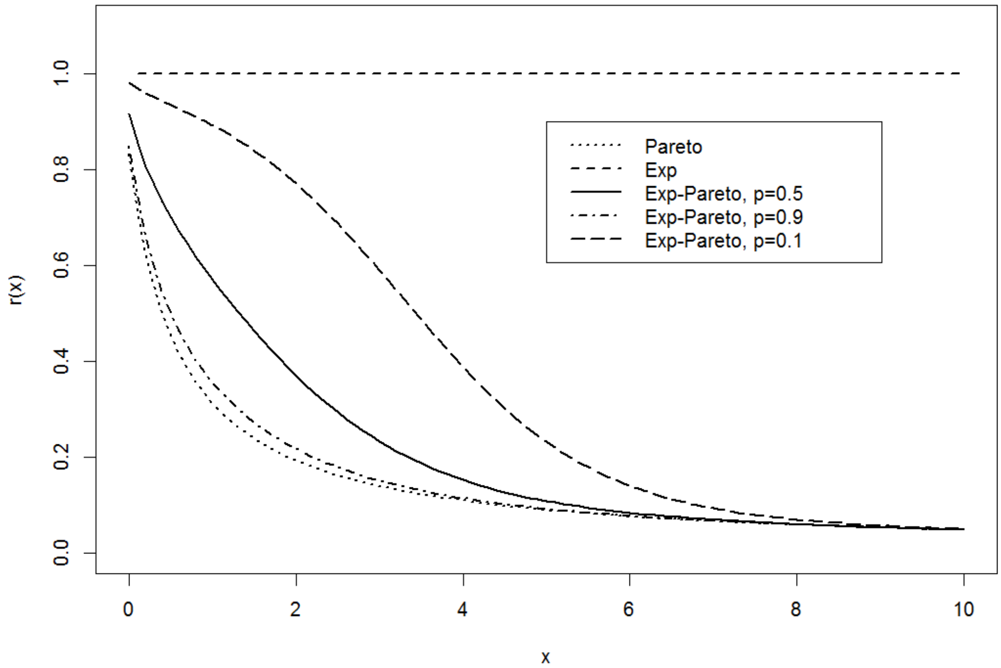

Figure 2: *The failure rate function of distributions* $Exp(1)$*,* $Pareto(0.5; 0.6)$ *and mixture* $Exp - Pareto(1; 0.5; 0.6; p)$ *with different mixing proportions* p*.*

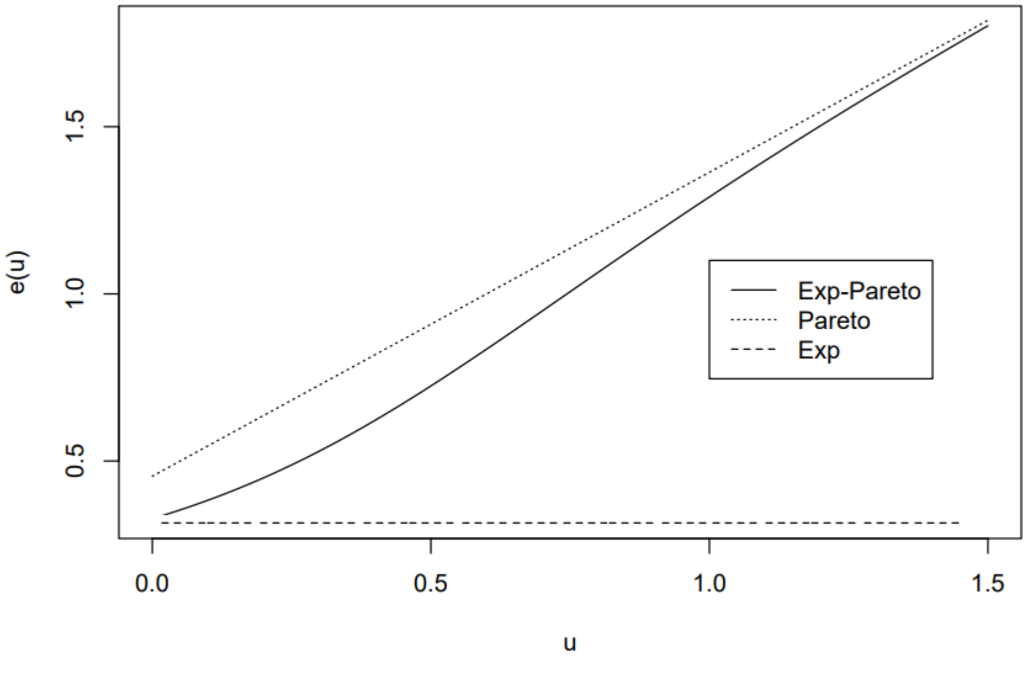

Figure 3: *Mean excess functions of distributions* $Exp(5)$*,* $Pareto(2.1; 0.5)$ *and mixture* $Exp - Pareto(5; 2.1; 0.5; 0.5)$*.*

The class of extreme value distributions (which combines all three types) is $G_\eta(cx + d)$ with real

$c > 0,\ d$, where

$$G_\eta(x) = \begin{cases} e^{-(1+\eta x)^{-1/\eta}}, & \eta \neq 0,\ 1+\eta x > 0; \\ e^{-e^{-x}}, & \eta = 0. \end{cases}$$

It is easy to check that the first component of Exponential-Pareto mixture distribution is in the maximum domain of attraction (MDA) of the Gumbel law, while the second is in MDA of the Frechet law.

For given $0 \leq \tau \leq \infty$ and a sequence $\{u_n,\ n \geq 1\}$ of real numbers the following are equivalent [10]

$$n\overline{F}(u_n) \to \tau \text{ as } n \to \infty \tag{21}$$

and

$$\mathbf{P}(M_n \leq u_n) \to e^{-\tau} \text{ as } n \to \infty. \tag{22}$$

If condition (20) is satisfied, then convergence (22) is preserved for any linear normalizing sequence $u_n(x) = a_n x + b_n,\ n \geq 1$ and expression (22) becomes

$$\mathbf{P}(M_n \leq u_n(x)) \to \tau(x),$$

where a concrete form of the function $\tau(x)$ depends on the type of the limiting distribution.

Theorem 2. Let the sequence of independent rv's $X_1, \ldots, X_n$ have Exponential-Pareto mixture distribution with df defined by (1). Define $M_n = \max(X_1, \ldots, X_n)$ is maxima among n values of sequence. Then $(M_n - b_n)/a_n \in MDA(\Phi_\alpha)$, where

$$a_n = x_0 n^{1/\alpha},\ b_n = -x_0. \tag{23}$$

Proof. First we find $\tau(x)$ substituting $u_n(x) = x_0 n^{1/\alpha}\, x - x_0$ into the relation (21):

$$\begin{aligned} n\overline{F}(u_n(x)) &= n\left[pe^{-\lambda u_n(x)} + \frac{(1-p)x_0^\alpha}{(x_0+u_n(x))^\alpha}\right] \\ &= n\left[pe^{-\lambda u_n(x)} + \frac{(1-p)x_0^\alpha}{(x_0\, n^{1/\alpha}\, x))^\alpha}\right] \\ &= \frac{n(1-p)x^{-\alpha}}{n}\left[\frac{p\, n\, e^{-\lambda u_n(x)} x^\alpha}{1-p} + 1\right] \\ &\longrightarrow (1-p)x^{-\alpha} \end{aligned}$$

since $u_n(x) \to \infty$ and $n\, e^{-\lambda u_n(x)} \to 0$ as $n \to \infty$. In accordance with (22) we get the following asymptotic distribution:

$$P(M_n \leq u_n(x)) \to e^{-(1-p)x^{-\alpha}} \text{ as } n \to \infty, \tag{24}$$

that is Frechet distribution. ∎

6. Uniform distance and Kullback-Leibler divergence

The uniform distance between two distributions F_X and F_Y [13],

$$\Delta(F_X, F_Y) = \sup_x |F_X(x) - F_Y(x)|, \tag{25}$$

is a widely used probability metric in sensitivity analysis [14].

It follows from the Pinsker-Csiszar Inequality [3] that uniform distance is bounded by Kullback-Leibler divergence, namely

$$\Delta(F_X, F_Y) \leq \sqrt{2D_{KL}(X||Y)}, \tag{26}$$

where

$$D_{KL}(X||Y) = \int_0^\infty f_X(x) \log\left(\frac{f_X(x)}{f_Y(x)}\right) dx \tag{27}$$

is Kullback-Leibler divergence.

It is shown in [3] that the minimum divergence between Exponential and Pareto distribution is reached at $\lambda = \frac{\alpha-1}{x_0}$ and

$$D_{KL}(Y||X)_{min} \leq \frac{1}{\alpha(\alpha-1)}. \tag{28}$$

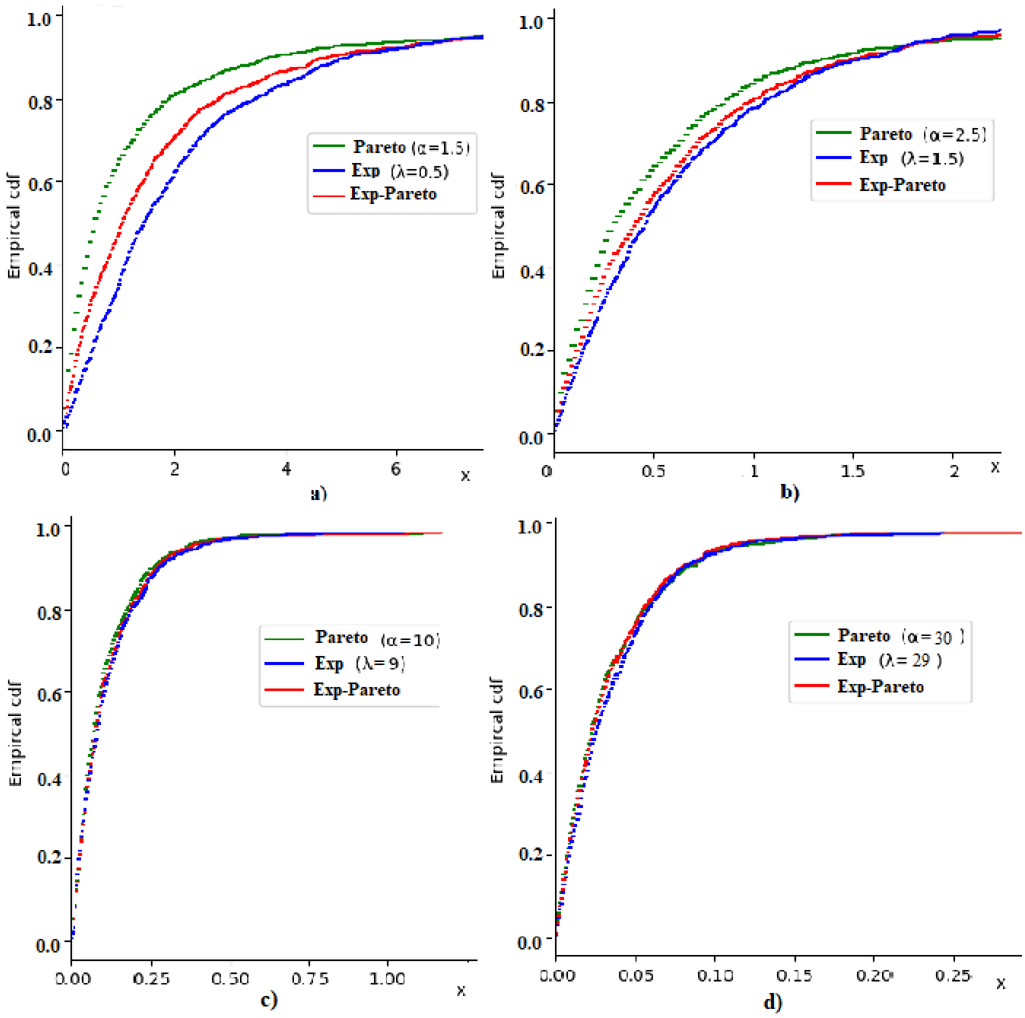

Figure 4: *Cumulative distribution functions of Exponential, Pareto and Exponential-Pareto distributions with different* α, λ, *and* $x_0 = 1, p = 0.5$

Clearly if the Kullback-Leibler divergence is close to zero, the uniform distance inherits this and thus implies that Exponential and Pareto distribution are close. We apply this to estimate the uniform distance between the Exponential-Pareto mixture and the Pareto distributions (for case $\alpha > 1$).

$$\begin{aligned}\Delta(F_Z, F_Y) &= \sup_x |pF_X(x) + (1-p)F_Y(x) - F_Y(x)| = \sup_x |pF_X(x) - pF_Y(x)| \\ &= p\Delta(F_X, F_Y) \le p\sqrt{2D_{KL}(Y||X)} \le \frac{\sqrt{2}p}{\sqrt{\alpha(\alpha-1)}}. \qquad (29)\end{aligned}$$

The last inequality demonstrates the convergence rate of $\Delta(F_Z, F_Y)$ to zero as $\alpha \to \infty$. We will get the same effect by approximating Pareto distribution via Exponential distribution as $\alpha \to \infty$ [3]. Let

$$\lambda = \frac{\alpha + o_\alpha(1)}{x_0}, \qquad (30)$$

where $o_\alpha(1) \to 0$ as $\alpha \to \infty$, then

$$\overline{F}_Y(x) = \lim_{\alpha\to\infty} \left(1 + \frac{\lambda x}{\alpha(1+o_\alpha(1))}\right)^{-\alpha} = e^{-\lambda x} = \overline{F}_X(x). \qquad (31)$$

The discussions above allow us to formulate the following lemma about approximation mixture by Exponential distribution for large α.

Lemma 5. Let df $F_Z(x)$ has form (1), then

$$F_Z(x) \to F_X(x) \;\text{ as }\; \alpha \to \infty, \;\text{ for all }\; x \ge 0,$$

where $F_X(x) = 1 - e^{-\lambda x},\; x \ge 0$.

Figure 4 demonstrates results of Exponential, Pareto and Exponential-Pareto mixture distributions with $x_0 = 1, p = 0.5$, a) $\alpha = 1.5, \lambda = 2.5$, b) $\alpha = 2.5, \lambda = 3.5$, c) $\alpha = 10, \lambda = 11$, d) $\alpha = 30, \lambda = 31$ for $n = 1000$ sample size. As expected, Kolmogorov-Smirnov test confirms that data sets from Exponential-Pareto mixture and Pareto distributions are homogeneous only for the case d) $\alpha = 30$. The uniform distance $\Delta(F_Z, F_Y)$ is $0.79, 0.34, 0.074, 0.012$, for cases a)-d), relatively. The upper bound for uniform distance $\frac{\sqrt{2}p}{\sqrt{\alpha(\alpha-1)}}$ defined by relation (29) is $0.81, 0.36, 0.074, 0.023$, relatively.

7. Parameters estimation

In this section we discuss the estimating the parameters of Exponential-Pareto mixture distribution by the method of moments and via maximization of log-likelihood function.

The method of moments gives the following estimates of the parameters α and λ, expressed in terms of the parameter x_0:

$$\begin{aligned}\alpha &= \frac{2(1-p)^2 x_0^2 - 2(1-p)x_0\overline{Z}}{S_Z^2 - (\overline{Z})^2} + 1; \\ \lambda &= \frac{2p(1-p)x_0 - 2p\overline{Z}}{2(1-p)x_0 - \overline{Z^2}},\end{aligned}$$

where $\overline{Z}$ – sample mean, $\overline{Z^2}$ – sample second moment, S_Z^2 – sample variance of random sample $x_!, \ldots, x_n$ from Exponential-Pareto mixture distribution. As x_0, it is possible to choose the first order statistic of random sample $x_!, \ldots, x_n$.

Let $x = (x_1, \ldots, x_k)$ be a realization of rv with Exponential-Pareto mixture distribution. Then likelihood function can be written as

$$L(x, \lambda, x_0, \alpha) = \prod_{k=1}^{n} (p\lambda e^{-\lambda x_k} + (1-p)\alpha x_0^\alpha (x_0 + x_k)^{-\alpha-1}).$$

The log likelihood function is given by

$$l(x, \lambda, x_0, \alpha) = \sum_{k=1}^{n} \log(p\lambda e^{-\lambda x_k} + (1-p)\alpha x_0^{\alpha}(x_0 + x_k)^{-\alpha-1}),$$

hence, the derivatives satisfy the following equations

$$\begin{aligned}
\frac{\partial l}{\partial \lambda} &= -p\lambda \sum_{k=1}^{n} \frac{x_k}{y_k} e^{-\lambda x_k}; \\
\frac{\partial l}{\partial x_0} &= (1-p)\alpha x_0^{\alpha-1} \sum_{k=1}^{n} \frac{x_0 + \alpha x_k}{y_k (x_0 + x_k)^{\alpha+2}}; \\
\frac{\partial l}{\partial \alpha} &= (1-p) x_0^{\alpha}(1 - \alpha \log(1 + 1/\alpha)) \sum_{k=1}^{n} \frac{1}{y_k (x_0 + x_k)^{1+\alpha}}; \\
\frac{\partial l}{\partial p} &= \lambda \sum_{k=1}^{n} \frac{e^{-\lambda x_k}}{y_k} - \alpha x_0^{\alpha} \sum_{k=1}^{n} \frac{1}{y_k (x_0 + x_k)^{1+\alpha}},
\end{aligned}$$

where $y_k = p\lambda e^{-\lambda x_k} + (1-p)\alpha x_0^{\alpha}(x_0 + x_k)^{-\alpha-1}$.

Setting the last equations equal to zero, the numerical maximum likelihood estimates of α, x_0, λ can be obtained by standard numerical methods like Newton Raphson method. The EM algorithm can be applied for iterative calculation of maximum likelihood estimates. Denote

$$g_{k1} = \frac{p_1 f_X(x_k|\lambda)}{p_1 f_X(x_k|\lambda) + p_2 f_Y(x_k|x_0, \alpha)}, \quad g_{k2} = \frac{p_2 f_Y(x_k|x_0, \alpha)}{p_1 f_X(x_k|\lambda) + p_2 f_Y(x_k|x_0, \alpha)},$$

where $p_1 = p, \; p_2 = 1 - p$. Then maximization by parameters of

$$\sum_{k=1}^{n} (g_{k1}(\log p_1 + \log f_X(x_k|\lambda)) + g_{k2}(\log p_2 + \log f_Y(x_k|x_0, \alpha)) \to \max$$

leads to the following relations

$$\begin{aligned}
p_j &= \sum_{k=1}^{n} g_{kj}/n, \quad j = 1, 2, \\
\lambda &= \frac{\sum_{k=1}^{n} g_{k1}}{\sum_{k=1}^{n} g_{k1} x_k}, \\
\alpha &= \frac{\sum_{k=1}^{n} g_{k2}}{\sum_{k=1}^{n} g_{k2} \log(z_k / x_0)},
\end{aligned}$$

where $z_k = x_k + x_0$ and parameter x_0 can be obtained from the equality

$$\frac{\sum_{k=1}^{n} g_{k2}}{\sum_{k=1}^{n} g_{k2}/z_k} = \frac{\sum_{k=1}^{n} g_{k2}}{\sum_{k=1}^{n} g_{k2} \log(z_k / x_0)}.$$

Table 1 demonstrates the results of identification of the distribution's parameters of the request processing time for web server 'dots.center' for different sample size of data. The web server processes industrial Internet data related to fuel consumption and operation of vessel equipment.

	Parameters of Exponential-Pareto distribution				
sample size	10^2	10^3	10^4	10^5	10^6
p	0.500	0.450	0.425	0.245	0.207
α	3.37	2.43	2.62	2.48	2.28
x_0	0.41	0.40	0.36	0.38	0.36
λ	6.45	6.71	6.67	6.70	6.47
MSE	0.070	0.037	0.017	0.240	0.370

Table 1: *The parameters of Exponential-Pareto distribution for the of the request processing time.*

8. Conclusion

In this paper we introduced the Exponential-Pareto mixture distribution. Characteristic function, failure rate function, mean excess, conditional excess distribution are derived. It is proved that this distribution belong to the subexponential distributions. Under condition on parameters $\lambda \geq \alpha/x_0$ Exponential, Exponential-Pareto mixture and Pareto distributions are ordered in failure rate, as well as conditional excesses and mean excesses. It is proved that the limiting distribution of maxima among n values of rv's with Exponential-Pareto distribution has Frechet-type form. The upper bound of uniform distance between Exponential-Pareto mixture and Pareto distributions is derived. The maximal likelihood estimates for parameters are given.

9. Funding

The publication has been prepared with the support of Russian Science Foundation according to the research project No. 21-71-10135, https://rscf.ru/en/project/21-71-10135/..

References

[1] Nanuwong N., Bodhisuwan W., Pudprommarat Ch. (2015). A New Mixture Pareto Distribution and Its Application. *Thailand Statistician* 13(2): 191-207.
[2] Marshall A., Olkin I. Life distributions. Structure of nonparametric, semiparametric and parametric families. Springer Science+Business Media, LLC., 2007.
[3] Weinberg G. V. (2016) Kullback-Leibler divergence and the Pareto-Exponential approximation. *SpringerPlus* 5(604)
[4] Nguyen-An H., Silverston T., Yamazaki T., Miyoshi T. (2021) IoT Traffic: Modeling and Measurement Experiments. *IoT* 2: 140–162.
[5] Peshkova, I., Morozov, E., Maltseva, M. (2020). On Comparison of Multiserver Systems with Exponential-Pareto Mixture Distribution. *Gaj, P., Guminski, W., Kwiecien, A. (eds) Computer Networks. CN 2020. Communications in Computer and Information Science, Springer, Cham* 1231.
[6] Peshkova I., Morozov E. (2022). On comparison of multiserver systems with multicomponent mixture distributions. *J. Math. Sci.* 267(2): 260–272.
[7] Peshkova, I., Golovin, A., Maltseva, M. (2023). On Waiting Time Maxima in Queues with Exponential-Pareto Service Times. *Vishnevskiy, V.M., Samouylov, K.E., Kozyrev, D.V. (eds) Distributed Computer and Communication Networks. DCCN 2022. Communications in Computer and Information Science Springer, Cham* 1748.
[8] Bulinski A., Slepov N. (2022). Sharp Estimates for Proximity of Geometric and Related Sums Distributions to Limit Laws. *Mathematics* 10(4747). https://doi.org/10.3390/math10244747

[9] Adler R. J., Feldman R. E. and Taqqu M. S. A Practical Guide to Heavy Tails: Statistical Techniques and Applications. Birkhauser Boston Inc., 1998.

[10] Embrechts P., Kluppelberg C., Mikosch T. Modelling Extremal Events for Insurance and Finance. Springer Heidelberg New York Dordrecht London, 1997.

[11] Navarro J. (2016). Stochastic comparisons of generalized mixtures and coherent systems *TEST*, 25: 150–169.

[12] Leadbetter M,, Lindgren G., Rootzen H. Extremes and related properties of random sequences and processes. Springer-Verlag New-York Inc. Springer series in statistics, 1983.

[13] Zolotarev V. M. Modern theory of summarizing of independent random variables. Nauka, Moscow, 1986. (in Russian)

[14] Korolev V. Yu, Krylov V. A. and Kuz'min V. Yu. (2011) Stability of finite mixtures of generalized Gamma-distributions with respect to disturbance of parameters *Inform. Primen.* 5(1): 31–38.