# 8. ON GENERALIZED PARETO DISTRIBUTIONS

**I. MIERLUS-MAZILU***

## Abstract

*This paper discusses a set of algorithms for numerical simulation of the Generalized Pareto Distribution and the propositions they are based on. The numerical results and the histogram of generating data are also presented.*

**Keywords:** Generalized Pareto Distribution; numerical simulation
**JEL classification**: C15, C16, C44

## 1. Introduction

The Generalized Pareto Distribution (GPD) was introduced by Pikands (1975) and has sine been further studied by Davison, Smith (1984), Castillo (1997, 2008) and other.

If we consider an unknown distribution function $F$ of a random variable $X$, we are interested in estimating the distribution function $F_u$ of variable of $x$ above a certain threshold $u$. The distribution function $F_u$ is so called the conditional excess distribution function and is defined as:

$$F_u(y) = P(X - u \le y \mid X > u),\ 0 \le y \le x_F - u,$$

where: $X$ is random variable, $u$ is a given threshold, $y = x - u$ are the excesses and $x_F$ is the right endpoint of $F$.

We verify that $F_u$ can be written in:

$$F_u(y) = \frac{F(u+y) - F(u)}{1 - F(u)} = \frac{F(x) - F(u)}{1 - F(u)}.$$

Pickands (1975) posed that for a large class of underlying distribution function $F$ the conditional excess distribution function $F_u(y)$, for $u$ large, is well approximated by:

$$F_u(y) \approx F(y, k, \sigma),\ u \to \infty$$

where :

* *Technical University of Civil Engineering, Bucharest, email mmi@mail.utcb.ro*

$$F(y;k,\sigma)=\begin{cases}1-\left(1-\dfrac{ky}{\sigma}\right)^{\frac{1}{k}} & k\neq 0,\ \sigma>0\\ 1-e^{-\frac{x}{\sigma}} & k=0,\ \sigma>0\end{cases}$$

for $y\in\left[0,(x_F-u)\right]$, if $k\leq 0$, and $y\in\left[0,\dfrac{\sigma}{k}\right]$, if $k>0$. $F(\cdot,k,\sigma)$ is the so-called Generalized Pareto Distribution. If $x$ is defined as $x=u+y$ the GPD can also be expressed as a function of $x$.

**Definition 1** Castillo (1997): *The random variable $X$ has a generalized Pareto distribution $GPD(k,\sigma)$ if the cumulative distribution function of $X$ is given by*

$$F(x;k,\sigma)=\begin{cases}1-\left(1-\dfrac{kx}{\sigma}\right)^{\frac{1}{k}} & k\neq 0,\ \sigma>0\\ 1-e^{-\frac{x}{\sigma}} & k=0,\ \sigma>0\end{cases}$$

*where: $\sigma$ and $k$ are scale and shape parameters, $x>0$ for $k\leq 0$ and $x\in\left[0,\dfrac{\sigma}{k}\right]$ for $k>0$.*

**Remark 1:**

- *When $k=0$, the $GPD(k,\sigma)$ reduces to the exponential distribution with mean's: $Exp(\sigma)$.*
- *When $k=1$, the $GPD(k,\sigma)$ becomes uniform : $U[0,\sigma]$.*
- *When $k<0$, the $GPD(k,\sigma)$ reduces to the $Pareto(k_1,a,c)$ distribution of the second kind.*

**Definition 2** Castillo (2008): *The random variable $X$ has the $Pareto(k_1,a,c)$ distribution of the second kind if the cumulative distribution function of $X$ is given by*

$$F(x;k_1,a,c)=1-\frac{k_1}{(x+c)^a}.$$

## 2. Algorithm for Numerical Simulation

### *2.1. The inverse method*

This method is based on the following lemma (Smirnov-Hincin).

**Lemma 1:** *Let $X$ be a random variable having $F$, the cumulative distribution function, inversable, and let $U$ be a uniform random variable on $(0,1)$. Then $Y = F^{-1}(U)$ has the same cumulative distribution function with $X$ (e. g. $Y$ is a sample of $X$ ).*

<u>Proof:</u> $P(Y < y) = P(F^{-1}(U) < y) = P(U < F(y)) = F(y), U$ being uniformly distributed on $(0,1)$ - q.e.d.

According to Lemma 1, we can write the following algorithm.

**Al.** *{Algorithm for numerical simulation of the $GPD(k,\sigma)$ random variable using the inverse method}*

$S_0$ Make suitable initializations; Input: $k, \sigma$ ;

$S_1$ if $k = 0$ then

repeat

Generate $U \to U(0,1)$; $X = -\frac{1}{\sigma}\ln U$

until stop condition

else

if $k = 1$ then

repeat

Generate $U \to U(0,1)$; $X := \sigma \cdot U$

until stop condition

else

input: $a = \frac{\sigma}{k}$;

repeat

Generate $U \to U(0,1)$; $X = a \cdot (1 - U^k)$

until stop condition

$S_2$ Stop

### *2.2. The method based on the transformation of the standard exponential variable*

We consider the case when $k > 0$ and $x \in \left(0, \frac{\sigma}{k}\right)$, according to Remark 1.

**Proposition 1:** Let $Y \to Exp(1)$. Then $X = \frac{\sigma}{k}\left(1 - e^{-kY}\right)$ is a $GPD(k,\sigma)$ random variable.

Proof:

$$F(x)=P(X<x)=P\left(\frac{\sigma}{k}\left(1-e^{-kY}\right)<x\right)=P\left(1-e^{-kY}<\frac{kx}{\sigma}\right)=P\left(1-\frac{kx}{\sigma}<e^{-kY}\right)=$$

$$=P\left(\ln\left(1-\frac{kx}{\sigma}\right)<-kY\right)=P\left(Y<-\frac{1}{k}\ln\left(1-\frac{kx}{\sigma}\right)\right)=P\left(Y<-\ln\left(1-\frac{kx}{\sigma}\right)^{\frac{1}{k}}\right)=$$

$$\xlongequal{Y\to Exp(1)} 1-e^{\ln\left(1-\frac{kx}{\sigma}\right)^{\frac{1}{k}}}=1-\left(1-\frac{kx}{\sigma}\right)^{\frac{1}{k}} \quad for \quad x\in\left(0,\frac{\sigma}{k}\right)$$

and $k>0$ - q.e.d.

**Remark 2:** *Since* $Y\to Exp(1)$, *with the inverse method we have* $Y=-\ln(U)$, $U\to U(0,1)$ *and thus we obtain:* $X=\frac{\sigma}{k}\left(1-e^{-k(-\ln U)}\right)=\frac{\sigma}{k}\left(1-e^{\ln U^k}\right)=\frac{\sigma}{k}\left(1-U^k\right)$.

Based on the proposition above, we have the following algorithm:

**AEx.** *{Algorithm for numerical simulation of the generalized Pareto distribution* $GPD(k,\sigma)$ *starting from a standard exponential random variable}*

$S_0$ Make suitable initializations; Input $k,\sigma$ ; Calculeazã $\alpha=\frac{\sigma}{k}$ ;

$S_1$ repeat

Generate $U\to U(0,1)$; $X:=\alpha\left(1-U^k\right)$

until stop condition

$S_2$ Stop

### *2.3. The mixture method*

We consider the case when $k<0$, $\sigma>0$ and $x>0$.

**Proposition 2**: Castillo (1997) *Let* $X_\theta\to G(x,\theta)=1-e^{\frac{\theta kx}{\sigma}}$ *with* $x>0$, $\sigma>0$, $k<0$ *and let* $\theta>0$, $\theta$ *a sample of the random variable* $\Theta\to Gamma\left(0,1,-\frac{1}{k}\right)$.

*Then the random variable* $X$ *obtained from mixing the two random variables* $X_\theta$ *and* $\Theta$ *is a* $GPD(k,\sigma)$ *random variable.*

<u>Proof:</u> $X_\theta\to 1-e^{\frac{\theta kx}{\sigma}}$, $\Theta\to\frac{1}{\Gamma(-\frac{1}{k})}\theta^{-\frac{1}{k}-1}e^{-\theta}$

We calculate the cumulative distribution function of the random variable $X$ :

$$F(x) = P(X < x) = \int_0^\infty \left(1 - e^{\frac{\theta kx}{\sigma}}\right) \frac{1}{\Gamma\left(-\frac{1}{k}\right)} e^{-\theta} \theta^{-\frac{1}{k}-1} d\theta =$$

$$= \frac{1}{\Gamma\left(-\frac{1}{k}\right)} \underbrace{\int_0^\infty e^{-\theta} \theta^{-\frac{1}{k}-1} d\theta}_{\Gamma\left(-\frac{1}{k}\right)} - \frac{\frac{1}{\Gamma\left(-\frac{1}{k}\right)} \int_0^\infty e^{-\theta\left(1-\frac{kx}{\sigma}\right)} \cdot \left[\theta\left(1-\frac{kx}{\sigma}\right)\right]^{-\frac{1}{k}-1} d\left[\theta\left(1-\frac{kx}{\sigma}\right)\right]}{\left(1-\frac{kx}{\sigma}\right)^{-\frac{1}{k}}} =$$

$$= 1 - \left(1-\frac{kx}{\sigma}\right)^{\frac{1}{k}} \cdot \frac{1}{\Gamma\left(-\frac{1}{k}\right)} \underbrace{\int_0^\infty e^{-\theta\left(1-\frac{kx}{\sigma}\right)} \cdot \left[\theta\left(1-\frac{kx}{\sigma}\right)\right]^{-\frac{1}{k}-1} d\left[\theta\left(1-\frac{kx}{\sigma}\right)\right]}_{\Gamma\left(-\frac{1}{k}\right)} =$$

$$= 1 - \left(1-\frac{kx}{\sigma}\right)^{\frac{1}{k}} \text{ - q.e.d.}$$

Based on the proposition above we can write the following algorithm:

**AMm.** *{Algorithm for numerical simulation of the* $GPD(k,\sigma)$ *distribution by the mixture method}*

S$_0$ Make suitable initializations; Input $k, \sigma, \alpha = \frac{1}{k}, \beta = \frac{\sigma}{k}$

S$_1$ repeat

Generate $\theta \to Gamma\left(0,1,-\frac{1}{k}\right)$

Generate $U \to U(0,1)$ {$U$ and $\theta$ are independent};

Deliver $X := \frac{\beta}{\theta} \ln U$

until stop condition;

S$_2$ Stop

The random variable $X_\theta$ is obtained by the inverse method, as before.

There are several methods for simulation of *Gamma* distribution. Here we will remind only some of them.

Therefore, the next problem is how we can generate $X$ a $Gamma(0,1,\nu)$ random variable.

### *2.3.1. The case when the parameter is greater than one*

*A rejection enveloping algorithm*

This method is based on the following theorem:

***Theorem 1*** *Văduva (1977): Let $Y$ be a random vector, to be generated, having the probability density function $f(y), y \in R^m$ and $Z$, another random vector, which can be generated, having the probability density function $h(z), z \in R^m$ such as $P(h(Z))=0)=0$. Assume that there is a positive and finite constant $\alpha$ such as:*

$$\frac{f(x)}{h(x)} \leq \alpha, \quad \forall x \in R^m, 1<\alpha<\infty$$

*If $U$ is an uniform$(0,1)$ random variate, independent from $Z$, then the conditional probability density function of $Z$, given that $0 \leq U \leq \frac{f(Z)}{\alpha h(Z)}$ is $f(y)$.*

The rejection procedure based on this theorem is sometimes called the enveloping procedure and the acceptance probability is $p_a = \frac{1}{a}$.

If in Theorem 2 we consider the case $m=1$, and that

$$f(y)=\frac{1}{\Gamma(\nu)} y^{\nu-1} e^{-y} I_{(0,\infty)}(y), \nu \geq 1 \text{ and } h(z)=\frac{k}{1+\frac{(z-(\nu-1))^2}{c}}, z \geq 0, c \geq 2\nu-1$$

meaning that $h(z)$ is the Nonstandard Cauchy Distribution, centered in $\nu-1$ with $c=2\nu-1$ and $K$ is a norming constant. It is shown that

$$a=\frac{1}{K\Gamma(\nu)}(\nu-1)^{\nu-1} e^{-(\nu-1)}, \; p_a=\frac{1}{a}$$

and hence the average number of trials (meaning that pairs of $(U,Z)$) for obtaining a sampling value of $Y$ is $e_n=\pi\sqrt{e}e^{-(\nu-1)}(\nu-1)^{\nu-1}\Gamma(\nu)$.

Note that $e_n \leq \sqrt{\pi}$ for $1-e<\nu<\infty$, which illustrates a good performance of the algorithm.

A disadvantage of this algorithm is that it uses three elementary functions $(\tan, e^x, \log)$ which may increase computing time and anyway will reduce accuracy. Therefore, Devroye (1986) have proposed a composition-rejection algorithm based on decomposition of the gamma density $f(y)$ in the form:

$$f(y)=p_1 f_1(y)+p_2 f_2(y), (p_1+p_2=1)$$
$$f_1(y)=0 \text{ if } y \notin [0,b]$$

$$f_2(y) = 0 \text{ if } y \notin (b, \infty)$$

where: $f_1(y)$ is enveloped by the truncated normal on $[0.b]$ and $f_2(y)$ is enveloped by the truncated exponential on $(0, \infty)$, $b$ being properly selected.

This is one of the best algorithms for generating a gamma random variable when $\nu > 1$.

*An algorithm based on the ratio-of-uniforms*

A new fast and more accurate algorithm for generating a standard gamma variate with the parameter $\nu > 1$ will be presented. This algorithm is based on the following:

**Theorem 2** Văduva (1977): If the probability density function $f$ of the random vector $X$ is in the form $f(x) = \frac{1}{H} h(x)$, $H = \int_{R^m} f(x)dx$, $d(m) = \frac{1}{mc+1}$, $\varphi : R^{m+1} \to R^m$ is in the form:

$$\varphi(v_0, v_1, ..., v_m) = \left( \frac{v_1}{v_0^c}, \frac{v_2}{v_0^c}, ..., \frac{v_m}{v_0^c} \right)', \quad c > 0$$

and $C = \left\{ (v_0, v_1, ..., v_m)' \mid \gamma(v_0, v_1, ..., v_m) \le 0 \right\}$ is bounded, and if $V$ is a random vector uniformly distributed on $C$, then the random vector $Y = \varphi(V)$ has the probability density function f.

In Theorem 2 we consider $m = 1$ and

$$f(x) = \frac{1}{H} h(x), H = \Gamma(\nu), h(x) = x^{\nu-1} e^{-x}$$

$$f(x) = \frac{1}{H} h(x), \quad H = \Gamma(\nu), \quad h(x) = x^{\nu-1} e^{-x}$$

$$\gamma(v_0, v_1) = \log(v_0) - \frac{1}{c+1} \left[ (\nu - 1) \log\left( \frac{v_1}{v_0^c} \right) - \frac{v_1}{v_0^c} \right]$$

then, using Theorem 2, after some calculations, we have

$$a_0 = 0, \quad b_0 = (\nu - 1)^{\frac{\nu-1}{c+1}} e^{-\frac{\nu-1}{c+1}}, \quad a_1 = 0, \quad b_1 = \left( \frac{c\nu + 1}{c} \right)^{\frac{c\nu+1}{c+1}} e^{-\frac{c\nu+1}{c+1}}$$

$$p_a(c) = \frac{\Gamma(\nu)}{(c+1)(\nu - 1)^{\frac{\nu-1}{c+1}} \left( \frac{c\nu + 1}{c+1} \right)^{\frac{c\nu+1}{c+1}} e^{-\nu}}$$

Now the final algorithm can be described easily.

An interesting question concerns the finding of the “best” constant $c > 0$ to maximize the acceptance probability $p_a(c)$. The analytical treatments of this problem necessitate complicated calculations. Therefore we designed a special procedure which, starting from the initial value $c_0 = 1$, searches (using a *small step* of variation of $c$,) a positive value $c$ in the neighborhood of this initial value, to give a greater probability. It results that $\lim_{\nu \to \infty} p_a(c) = 0$, which shows that the algorithm has a bad behavior for large values of $\nu$. However, for not very large values of $\nu$ (e.g. $\nu \le 34$) it is faster than the algorithm based on a Cauchy envelope, as computer tests have shown.

### *2.3.2. The case when the parameter* $0 < \lambda < 1$

*A rejection algorithm*

This algorithm uses theorem 1 where

$$f(y) = \frac{1}{\Gamma(\nu)} y^{\nu-1} e^{-y} I_{(0,\infty)}(y), \quad 0 < \nu < 1, \quad h(z) = \nu z^{\nu-1} e^{-z} I_{(0,\infty)}(z)$$

(meaning that $h$ is a standard Weibull probability density function) which gives

$$\frac{f(y)}{h(y)} \le a(\nu) = \frac{1}{\Gamma(\nu+1)} e^{\xi(1-\nu)}, \quad \xi = \nu^{\frac{\nu}{1-\nu}}, \quad p_a = \frac{1}{a}$$

In our application we have to take into account that the behavior of the *rejection* algorithm is good mainly for $0.12 \le \nu \le 0.9999$ as computer tests have shown (Văduva, 1977).

## 3. Numerical Results

I this paper, we implement all these three algorithms: AI{Algorithm for numerical simulation of the $GPD(k,\sigma)$ random variable using the inverse method}, AEx{Algorithm for numerical simulation of the generalized Pareto distribution $GPD(k,\sigma)$ starting from a standard exponential random variable}, AMn{Algorithm for numerical simulation of the $GPD(k,\sigma)$ distribution by the mixture method}. Numerical application was performed using Mathcad.

```
GPD(k, σ, N) := | for i ∈ 1.. N          if k = 0
                |    X_i ← −ln(rnd(1)) / σ
                | for i ∈ 1.. N       if k = 1
                |    X_i ← σ·rnd(1)
                | otherwise
                |    | a ← σ / k
                |    | for i ∈ 1.. N
                |    |    X_i ← a·(1 − rnd(1)^k)
                | return X
```

The general algorithm *AI*

```
GPD(k, σ, N) := | a ← σ / k
                | for i ∈ 1.. N
                |    X_i ← a·(1 − rnd(1)^k)
                | return X
```

The general algorithm *AEx*

The general algorithm *AI* was called 1000 times. We consider $k = 2, \sigma = 2.5$ in Figure 1 and $k = 0, \sigma = 2.5$ in Figure 2.

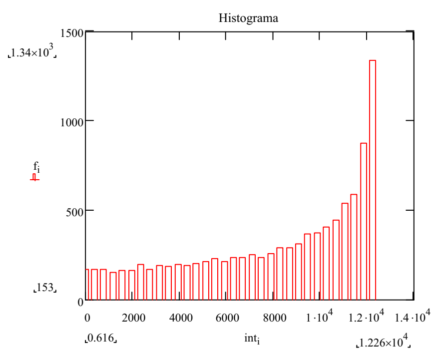


Fig. 1.

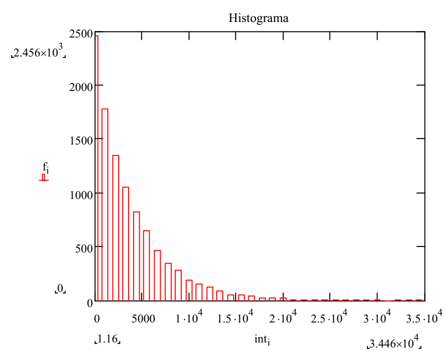


Fig. 2

The general algorithm *AEx* was called 1000 times. We consider $k = 2, \sigma = 2.5$ in Fig 3. Also in Fig 4 we use *AMn* algorithm called 1000 times, with $k = 2, \sigma = 2.5$.

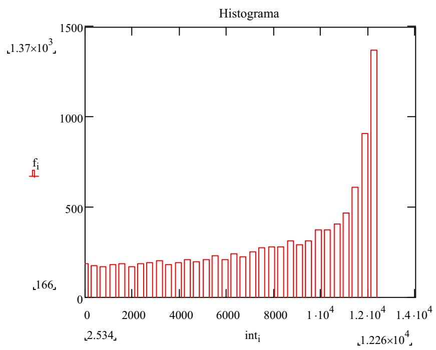


Fig. 3.

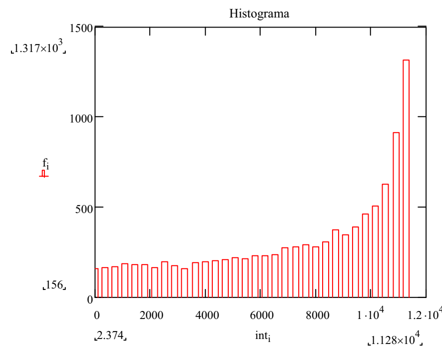


Fig. 4

The histogram obtained on generated selection shows that the empirical distribution is similar to the theoretical distribution.

## 4. Example and Conclusions

The GPD has applications in a number of fields, including reliability studies, in the modelling of large insurance claims, as a failure time distribution. Also it plays an important role in modeling extreme events. A model is frequently used in the study of income distribution and in the analysis of extreme events, e.g. for the analysis of the precipitation data, in the flood frequency analysis, in the analysis of the data of greatest wave heights or sea levels, maximum winds loads on buildings, in the maximum rain fall analysis, in the analysis of greatest values of yearly floods, breaking strength of materials, aircraft loads, etc.

The GPD has been quite popular not only for flood frequency analysis but for fitting the distribution of extreme natural events in general.

Every portfolio of risk policies incurs losses of greater and lower amounts at irregular intervals. The sum of all the losses (paid & reserved) in any one-year period is described as the annual loss burden or simply total incurred claims amount. An insurance company generally decides to transfer the high cost of contingent capital to a third party, a reinsurance company.

It is possible to generate randomly a number of claims per year and to calculate each claim severity through the Generalized Pareto distribution. In this case if we calculate the economic risk capital defined as the difference between the expected loss, defined as the expected annual claims amount, and in this case the 99.93th quantile of the distribution corresponding to a Standard & Poor's A rating.

The analysis highlights the importance of a reinsurance program in terms of "capital absorption & release" because what happens in the tail of the loss distribution - where things can go very wrong and where the inevitable could sadly happen - has relevant impact on the capital base. This implicitly demonstrates that reinsurance decisions based on costs comparison could lead to wrong conclusions.

## References


Castillo, E. and A.S. Hadi, (1997), "Fitting the Generalised Pareto Distribution to Data", *JASA*, 92( 440): 1609 – 1620.

Castillo J., Daoudib J. (2008), "Estimation of the generalized Pareto distribution", *Statistics & Probability Letters*, In Press.

Corradin, S. and B. Verbrigghe, (2003), "Economic Risk Capital and Reinsurance: an Application to Fire Claims of an Insurance Company" ASTIN Colloquium International Actuarial Association - Brussels, Belgium.

Davidson A.C. (1984), "Modeling excesses over high threshold with an application", In: *J. Tiago de Oliveria (ed.), Statistical Extremes and Applications*, Reidel, Dordrech, pp. 416-482.

Devroye L. (1986), "*Non-Uniform Random Variate Generation*", New York - Berlin Heidelberg - Tokyo: Springer Verlag,.

Deak I. "*Random Number Generators and Simulation*", Akademia Kiado, Budapest, 1990.

Johnson N., Kotz S. (1972), "*Distributions in statistics. Continous univariate distributions*", New York : John Wiley and Sons.

Pickands J. (1975), "Statistical Inference Using Extreme Order Statistics," *The Annals of Statistics*, 3: 119-131.

Vaduva I. (1997), "*Simulation Models for Computer Applications*" (in romanian), Bucharest : Technical Publishers.