Proceedings of the 8th International Conference on Applied Informatics
Eger, Hungary, January 27–30, 2010. Vol. 1. pp. 239–246.

Parameter Estimation and Random Number Generation to Stable Distributions

Csilla Csendes

Department of Applied Mathematics, University of Miskolc
e-mail: matcs@uni-miskolc.hu

Abstract

In this paper a robust parameter estimation based on the method of moments called Probability Integral Transformation is represented. The method is used to determine the characteristic exponent α, the scale γ and location parameter δ of a univariate symmetric stable distribution simultaneously from a random sample. A simulation sequence is made to test the accuracy and robustness of the estimation. In the second part we describe algorithms to multivariate stable random number generation in case of a discrete spectral measure and in independent case using Zolotarev's formula. The simulation results are illustrated.

Keywords: characteristic exponent α, stable distributions, random number generation

MSC: 60E07, 62F35, 62G35

1. Introduction

A wide research work was made on characterization of univariate stable distributions. Reliable methods were created to estimate the parameters and compute the density, however we still don't know analytical formulas for the density and cumulative distribution function (c.d.f) in general. The Probability Integral Transformation method described in Section 2 is a robust technique and more accurate then previously investigated algorithms which used the semi-interquartile range of the sample to evaluate the parameters. [4, 5]

Defining and describing multivariate stable distributions are more difficult, because the dependence structure appears between the components. Considering a general stable distribution the so-called spectral measure is used to evaluate the dependencies. It's hard to estimate a multivariate c.d.f because of computational time of high dimensional integrals.

Random number generation is the starting point to develop estimators to multivariate stable laws. In Section 3 some known methods applied to stable variables are summarized and compared in terms of their usability in practice.

2. Probability Integral Transformation

2.1. PT estimator for characteristic exponent α

PT estimation [6] is used to determine the characteristic exponent α of a univariate stable law by using the known probability density functions (p.d.f.) of normal and Cauchy distributions. We assume that $1 < \alpha < 2$ and calculate the scale (and location) parameter of the distribution represented by a sample using these two known density. The scale parameter values can be computed with an iterative algorithm considering a fixed α. If we change this α and simultaneously use the normal and Cauchy p.d.f's then we get scale parameter estimations equal at the sample's real α parameter.

If the distribution function F of ξ random variable can be inverted then $F(\xi)$ is uniformly distributed on $(0, 1]$. Let $F(x) = F_0(\frac{x-\mu}{\sigma})$, F and F_0 has the same type, and we define location μ and scale parameter σ according to F_0 standard distribution function. For uniform distributions:

$$E_F(F_0(\frac{\xi-\mu}{\sigma})) = \frac{1}{2}, \tag{2.1}$$

$$D_F^2(F_0(\frac{\xi-\mu}{\sigma})) = \frac{1}{12}, \tag{2.2}$$

The equation system with μ and σ unknown variables can be solved with an iterative algorithm called *ping-pong method*. We have 2.3,2.4 for the scale and location parameters. We take them in turns and repeat computation until reaching arbitrary precision.
Location parameter:

$$T_n^{(m+1)} = T_n^{(m)} + \frac{s\sum_{i=1}^{n}\psi(\frac{\xi_i - T_n^{(m)}}{s})}{n}, \tag{2.3}$$

Scale parameter:

$$[s_n^{(m+1)}]^2 = \frac{1}{(n-1)\beta}\sum_{i=1}^{n}\psi^2(\frac{\xi_i - T_n^{(m+1)}}{s_n^{(m)}})[s_n^{(m)}]^2, \tag{2.4}$$

where the initial values are

$$T_n^{(0)} = med\{\xi_i\}, \quad \text{and} \qquad s_n^{(0)} = C \cdot MAD. \tag{2.5}$$

and

$$\psi(x) = F_0(x) - \frac{1}{2}. \tag{2.6}$$

MAD denote the median absolute deviation.

$$MAD(\xi) = med(|x_i - med(\xi)|), \tag{2.7}$$

where $med(\xi)$ is the median of the distribution, s is the current estimation of the scale parameter. The constant C is used to have the initial estimation unbiased.

β is a function in equation (2.4) which is $\frac{1}{12}$ if the sample has the same type as F_0 and

$$\beta = D_G^2(\psi(\xi)), \tag{2.8}$$

otherwise. G denote ξ's distribution function.

In interval $(1, 2)$ we have to approximate $\beta(\alpha)$ function, because β is given in integral form (see [6]), but numerical integration has very high computational demand. At points $\alpha = 1, \alpha = 1.1, \alpha = 1.2, ...\alpha = 2$ a sample with 20 million elements were generated and $\beta(\alpha)$ values were determined for each case of normal and Cauchy distribution. Table 1. shows the results.

α	β_1	β_2
1	0.0833309645949110	0.126807877965645
1.1	0.0758844534723818	0.118966259082521
1.2	0.0697612892957584	0.112284032323310
1.3	0.0646999988570841	0.106570898511029
1.4	0.0604648399039825	0.101682622480835
1.5	0.0569093151515006	0.097443890906153
1.6	0.0538933607717261	0.093798682214659
1.7	0.0513226066667932	0.090637100836610
1.8	0.0491126022363082	0.087875629036068
1.9	0.0472087085432832	0.085445768785679
2	0.0455654051822800	0.083333333333333

Table 1: Values for β_1 and β_2 depend on α.

For increase the speed of estimation $\beta(\alpha)$ are approximated with a rational function

$$\beta_i(x) = \frac{a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0}{x^4 + b_3x^3 + b_2x^2 + b_1x + b_0}, \tag{2.9}$$

where $i = 1, 2$ and the coefficients $a_5, a_4, a_3, a_2, a_1, a_0, b_3, b_2, b_1, b_0$ were calculated via values in Table 1. The equation system was solved with MATLAB at first dropping the equation in case of $\alpha = 1.9$ after $\alpha = 2$ and then solved the whole system. The solution which had the smallest least square differences from values in Table 1. was decided to be the best. Table 2. shows the best coefficients.

By changing α, scale parameter estimations with $\beta_1(\alpha)$ and $\beta_2(\alpha)$ represent two monotonically increasing curves if we illustrate α on horizontal axis and scale parameter on vertical axis. The only one intersection point is at real α of the sample's distribution. If real α is near the end of the interval it's possible that the

coefficient	β_1	β_2
b_3	-3.83008202167381	-5.44424585925350
b_2	4.78393407388667	8.41128641608921
b_1	-2.07519730244991	-0.91519048820337
b_0	0.18011293964047	-4.11676503125739
a_5	0.00557315701358	0.02536047583564
a_4	-0.02655295929925	-0.20581738159343
a_3	0.12169846973572	0.84848196461615
a_2	-0.31598423581221	-2.18605774383017
a_1	0.34060543137722	3.06692780009580
a_0	-0.12044269481921	-1.68393473522529

Table 2: Coefficients for the rational fraction function.

no.	α	sample elements	repetition
1	1.3	100	100
2	1.5	100	100
3	1.8	100	100
4	1.5	100	2500
5	1.5	2500	100

Table 3: Simulations testing the PT estimation.

intersection point falls out $(1, 2)$ because of the random effect and estimation inaccuracy. The occurrence of this problem can be reduced by increasing the sample. Intersection point is founded by cut-and-try method on interval (1,2).

2.2. Statistical analysis of PT estimator for α

PT estimation was tested with Monte-Carlo method. The results were analyzed with *Statistica for Windows 8.0*. Table 3. contains simulation details. Theoretical assumption is that the estimated α values follow normal distribution with mean of the generated random sample. The hypothesis was tested with chi-squared test.

When the sample had 100 elements (in 1^{st}-4^{th} simulations) the procedure gave wrong values, that's why in probability tests less case performs than the ideal.

In the 1^{st}-3^{rd} simulations 100 repetition was made. Here we can accept the normality in 95% significance level. In the 1^{st} and 3^{rd} simulation a few estimations were fallen outside, but none in the 2^{nd}. In the 4^{th} simulation I had 2500 repetitions and 6 fallen-out values.

In the 4^{th} simulation we have to reject the normality. The mean has the same difference from original α than in the 1^{st}-3^{rd} simulations, but the standard deviation increased. Estimations were spread on the whole interval (and out) so a

sample with 100 elements is too less to have a reliable estimation.

In the 5^{th} on interval $[1.4, 1.6]$ chi-squared test gave very high p-value proving that the estimations follow the normal distribution and no bad value popped up. Obviously the precision could be better if we raise the sample. But it's not possible boundlessly because of the high computation time. Samples at the order of this magnitude in simulation 5^{th} give good estimations.

Table 4. shows confidence intervals for α in case of different simulations. α was set to 1.5 to avoid wrong cases. *Lower* and *upper* means the confidence interval's bounds.

Nr.	elem.	repet.	mean	std.dev.	lower	upper
1	50	100	1,5278	0,2176	1,4841	1,5716
2	50	2500	1,5322	0,2099	1,5237	1,5408
3	50	10000	1,5266	0,2103	1,5223	1,5309
4	100	100	1,5133	0,1748	1,4783	1,5482
5	100	2500	1,5169	0,1552	1,5107	1,5231
6	100	10000	1,5212	0,1556	1,5180	1,5243
7	400	100	1,5106	0,0897	1,4926	1,5285
8	400	2500	1,5073	0,0770	1,5042	1,5104
9	400	10000	1,5073	0,0777	1,5057	1,5088

Table 4: Confidence intervals for α

3. Random number generation to multivariate stable distributions

3.1. Universal generators

Devroye's [2] survey to nonuniform random variate generation provides us classical paradigms such as the inversion, rejection, and alternating series method.

For a univariate random variable the inversion method is the following. Assume that we have the distribution function F and F^{-1} inverse of F. Generate a random variable X as $F^{-1}(U)$ where U is uniform on $[0, 1]$, then X is distributed according to F. This method is chosen when F and F^{-1} are computable quickly, but unfortunately in the case of stable distributions they aren't.

Another way is to transform a uniform random variable with standard routines, for example trigonometric, exponential or logarithmic functions. Although stable distributions cannot be transformed from one uniform variable, they can from two. Using integral representation of the c.d.f for general α and β Chambers et al. [1] gave a solution. Zolotarev [8] has similar formulas using a uniform U and an exponential E variable.

If calculation of F^{-1} is impossible we can use the acceptance-rejection method assuming we know the density function $f(x)$ to F. The basic idea is instead generating with F we try to find a distribution G with density $g(x)$ which is as close to $f(x)$ as possible and can be generated quickly. The ratio $c = sup\{f(x)/g(x)\}$ is the rejection constant. The algorithm's steps are: generate a random variable Y with G and a uniform random variable U. If $U \leq \frac{f(Y)}{cg(Y)}$ then we "accept" Y, so set $X = Y$. Otherwise reject it, so generate another Y. The method's time complexity is proportionate to constant c. For our case it's hard to find appropriate dominating curves to stable density because we don't even know it exactly.

The alternating series method developed by Devroye [3] can be used when f is known as infinite series. It suffices to have an approximation $\phi_n(x)$ that tends to $f(x)/(cg(x))$ as $n \to \infty$ and for which we know a monotone error bound $\epsilon_n(x) \to 0$. In that case we let n increase until for the first time, either

$$U \leq \phi_n(X) - \epsilon_n(X) \qquad \text{(in which case we accept X) or} \tag{3.1}$$

$$U \geq \phi_n(X) + \epsilon_n(X) \qquad \text{(in which case we reject X).} \tag{3.2}$$

3.2. Independent case

While we have a general formula to generate a one dimensional stable variable it's obvious to generate independently the components and piece together them in a vector. This way an arbitrary dimensional variable can be generated. The well-known formula from Zolotarev [8] is used to made three dimensional vectors.

Figure 1. and Figure 2. show surfaces approximated from relative frequency values with $\alpha = 1.8$ and $\alpha = 1.5$ (sample size :100000), Figure 3. and 4. show the scattered sample respectively.

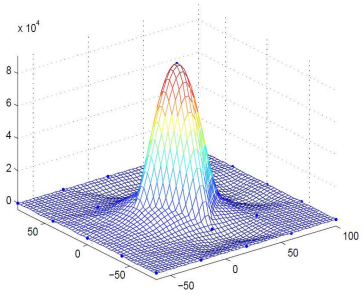

Figure 1.

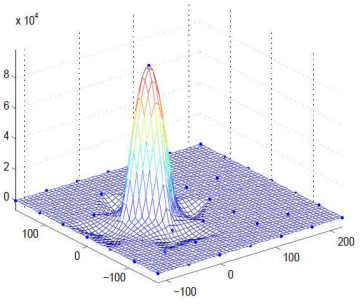

Figure 2.

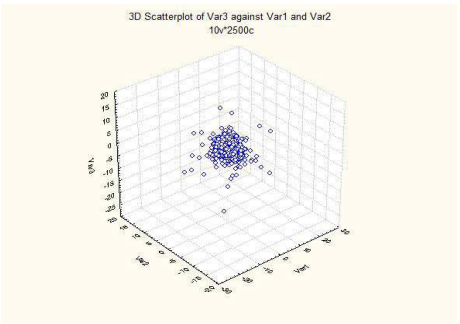

Figure 3.

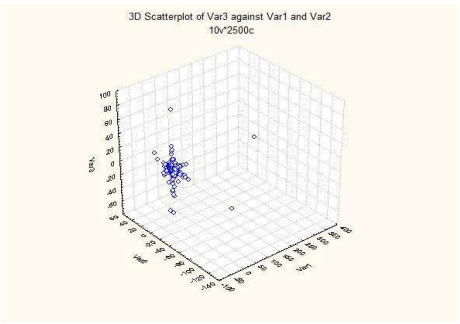

Figure 4.

3.3. Case of discrete spectral measure

This procedure coincident with the case of using a discrete spectral measure presented by J. P. Nolan [7]. Consider a discrete spectral measure Γ with a finite number of point masses

$$\Gamma(.) = \sum_{j=1}^{n} \gamma_j \delta_{s_j}(.) \tag{3.3}$$

where γ_j's are the weights and δ_{n_j}'s are point masses at the points $s_j \in S_d$, $j = 1, ...n$. The characteristic function of $X \sim S_{\alpha,d}(\Gamma, \mu_0)$ is

$$\phi_X(t) = Eexp\{i <X, t>\} = exp(-I_X(t) + i <\mu_0, t>) \tag{3.4}$$

where the function in the exponent is

$$I_X(t) = \int_{S_d} \phi_\alpha(<t, s>)\Gamma(ds) \tag{3.5}$$

and

$$\phi_\alpha(u) = \begin{cases} |u|^\alpha(1 - isign(u)\tan(\pi\alpha/2)), & \text{if } \alpha \neq 1, \\ |u|(1 + i(2/\pi)sign(u)\ln|u|), & \text{if } \alpha = 1. \end{cases}$$

For a discrete spectral measure the characteristic function takes the form

$$\phi^*(t) = exp\Big(-\sum_{j=1}^{n} \psi_\alpha(<t, s_j> \gamma_j)\Big) \tag{3.6}$$

This expression is numerically simple while it's difficult to compute $\phi(t)$. Nolan's theorem says that for a given $\epsilon > 0$ there exist an $n = n(d, \alpha, \epsilon, \Gamma)$ and values $s_1, ..., s_n$ and $\gamma_1, ...\gamma_n$ so that

$$sup|p(x) - p^*(x)| < \epsilon \tag{3.7}$$

and p and p^* denote the density corresponding to the characteristic functions (3.4), (3.6). The approximation is useful for calculating $p(x)$ pointwise.

Definition 3.1. For a set $H \subset S_d$ define the cone generated by H as

$$Cone(H) = \{x \in \mathbb{R}^d . |x| > 0, x/|x| \in H\} = \{ra : r > 0, a \in H\}.$$

Theorem 3.2 (Araujo and Gine,1980).

$$\lim_{r\to\infty} \frac{P(X \in Cone(H), |X| > r)}{P(|X| > r)} = \frac{\Gamma(H)}{\Gamma(S_d)}$$

In words, the mass that Γ assigns to H determines the tail behavior of X in the 'direction' of H. Applied to a discrete spectral measure Theorem 1. says that the density will have creases along the rays starting at the origin and passing through the point masses. Graphically, the level curves of a two dimensional p.d.f. will be star-shaped. See more examples in [7].

Figure 1. and Figure 2. shows the same consequences with the points $(1,0,0)$, $(0,1,0)$, $(0,0,1), (-1,0,0)$, $(0,-1,0), (0,0,-1)$. The estimation with arbitrary points and point masses based on the results of Modarres and Nolan (1994). If X has characteristic function (3.6) then

$$X \stackrel{D}{=} \begin{cases} \sum_{j=1}^{n} \gamma_j^{1/\alpha} Z_j s_j, & \alpha \neq 1, \\ \sum_{j=1}^{n} \gamma_j (Z_j + \frac{2}{\pi} \ln \gamma_j) s_j, & \alpha = 1, \end{cases}$$

where $Z_i \sim S_\alpha(1,1,0)$.

References

[1] Chambers, J. M., Mallows, C. L.,Stuck, B. W.: A method for simulating stable random variables, *J. Amer. Statist. Ass.*, Vol. 71 (1976), 340–343.

[2] Devroye, L. *Non-Uniform Random Variate Generation*, New York: Springer (1986).

[3] Devroye, L. The series method in random variate generation and its application to the Kolmogorov-Smirnov distribution, *American Journal of Mathematical and Management Sciences*, Vol. 1, (1981), 359–379.

[4] Fama, E.F., Roll, R. Some Properties of Symmetric Stable Distributions, *Journal of the American Statistical Association*, Vol. 63 (1968), 817–836.

[5] Fama, E.F., Roll, R. Parameter Estimates for Symmetric Stable Distributions, manuscript, University of Chicago, (1969).

[6] Fegyverneki, S. Robust estimators and probability integral transformation, *Math. Comput. Modelling*, Vol. 38 (2003), 803–814.

[7] Nolan, J. P., Multivariate stable distributions: approximation, estimation, simulation and identification. In R. J. Adler, R. E. Feldman, and M. S. Taqqu (Eds.), A Practical Guide to Heavy Tails, 509–526. Boston: Birkhauser (1998).

[8] Zolotarev, V. M.: *One-Dimensional Stable Distributions*, American Mathematical Society, Providence, (1986).

Csilla Csendes
University of Miskolc, Miskolc-Egyetemváros 3515 Hungary