THE ASYMPTOTICS OF MOMENTS FOR THE REMAINING TIME OF THE HEAVY-TAIL DISTRIBUTIONS

Vladimir Rusev, Alexander Skorikov

National University of Oil and Gas "Gubkin University", Moscow, Russian Federation;

rusev.v@gubkin.ru, skorikov.a@gubkin.ru

Table of Contents

Heavy-tailed

This paper falls into category exploring the classical models with heavy tails: Gnedenko-Weibull, Benktander,Burr distributions. The moment‘s asymptotics for residual time have been derived especially for this heavy-tailed distributions.

Definition 1. One defines a distribution function (DF) F to be (right-) heavy-tailed if

$$\lim_{x \to +\infty} e^{\lambda x} \bar{F}(x) = \infty$$

for the tail of the DF F: $\bar{F} = 1 - F$ and for all $\lambda > 0$.

Definition 2. A distribution function F is called light-tailed if and only if it fails to be heavy-tailed.

The slow decrease the tail of distribution leads to the fact that random variable can take on very large values with a positive probability

Go to graphs

Mathematical models and Heavy-tailed

Recent mathematical models of computer systems and telecommunications networks are based on heavy-tailed distributions. For example the superposition of ON/OFF processes with a heavy-tailed distribution was presented by W. Willinger, M.Taqqu, R.Sherman and D.V. Wilson, [**Will**]. Such processes converges to a self-similar (fractal) process with self-similarity Hurst exponent $H = (3 - \alpha)/2$, where $0 < \alpha < 2$ is the tail index (or coefficient of heavy - tailedness).

[Will] Willinger W., Taqqu M. S., Sherman R. and Wilson D.V.,*Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level* (*IEEE/ACM Transactions on Networking*, **1997**, *Vol.5, No. 1*, 71–86.).

Ethernet traffic

It is known a correspondence between the generated traffic and the actual Ethernet traffic. This self-similar processes remain unchanged at varying time scales unlike Poisson processes, which flatten out as time scales change.

Go to graphs

On the other hand, the accuracy of the simulation obtained depends on the accuracy of the approximation of the distributions with a heavy tail used.

Residual Life Time

Let X be non-negative random variable with cumulative distribution function F. Consider the random variable $X_t = (X - t\,|X > t)$, which is called in survival analysis the residual life time with distribution function F_t or distribution of the excess over a threshold t.

The Mean Excess Function and Excess Variance

The mathematical expectation of random variable X_t is the mean residual life (MRL) or mean excess function (ME). The ME term is using in risk management, actuarial science and other extreme value problems. This function is defined as

$$\mu(t) = E(X - t\,|X > t) = \int_t^{+\infty} \overline{F}(x)dx/\overline{F}(t) \quad end \tag{1}$$

$$\sigma^2(t) = E(X_t^2) - \mu^2(t) = \frac{2}{\overline{F}(t)} \int_t^{+\infty} \overline{F}(x) \cdot \mu(x)\, dx - \mu^2(t). \tag{2}$$

is residual variance (excess variance).

Gnedenko-Weibull Distribution (GWD)

1. Gnedenko-Weibull
   The tail two-parameter Gnedenko-Weibull distribution
   $$\overline{F}(t;\alpha,\beta) = e^{-(\alpha t)^{\beta}}, \quad t \geq 0$$
   is a heavy tail if $0 < \alpha, \quad 0 < \beta < 1$.
2. Let T be a random variable with the two-parameter Gnedenko-Weibull distribution, then $\forall\alpha > 0, \ \forall\beta > 0$ holds

Mean Excess Function (GWD)

$$\mu(t) = \frac{t}{\beta(\alpha t)^{\beta}}\left(1 + \frac{1-\beta}{\beta(\alpha t)^{\beta}} + \frac{(1-\beta)(1-2\beta)}{\beta^2(\alpha t)^{2\beta}} + O\left(\frac{1}{t^{3\beta}}\right)\right), as \quad t \to \infty, \qquad (3)$$

The first term of expansion (3) is known, [**Em**] and complete formula (3) is original, [**RS**].

[Em] Embrechts P., Klüppelberg C., Mikosch T.*Modelling Extremal Events for Insurance and Finance.*; Publisher: Berlin, Springer, 1997.

[RS] Rusev, V.; Skorikov, A.Springer Proceedings in Mathematics and Statistics. **2021** *Vol.371*, 292–305

Excess Variance Function

Excess Variance Function see in [**RS**].

$$\sigma^2(t) = \frac{1}{\beta^2\alpha^{2\beta}t^{2(\beta-1)}}\left(1 + \frac{4(1-\beta)}{\beta(\alpha t)^{\beta}} + \frac{(1-\beta)(11-17\beta)}{\beta^2(\alpha t)^{2\beta}} + O\left(\frac{1}{t^{3\beta}}\right)\right), t \to \infty. \qquad (4)$$

Benktander type I distribution

1. Definition.
Benktander type I distribution is close to lognormal distribution. This distributions are using to model heavy tailed losses in actuarial science. Theoretically predicted sizes by Benktander I distribution are consistent with statistical size distributions in economics and actuarial sciences. The Benktander type I distribution function tail is given by

$$1 - F(x) = \left(1 + 2\frac{\beta}{\alpha}\ln x\right) e^{-\beta(\ln x)^2 - (\alpha+1)\ln x}, \quad \alpha > 0,\ \beta > 0,\ x > 0.$$

Benktander type I distribution. Moments

1. Type I Benkthander distribution has the following exact equality for the mean exceedance, [**Em**]:

$$\mu(t) = \frac{t}{\alpha + 2\beta \ln t}. \tag{5}$$

2. We derive residual variance:

$$\sigma^2(t) = \frac{t^2}{(\alpha + 2\beta \ln t)^2} \left[1 + \frac{2}{2\beta \ln t + \alpha - 1} + O\left(\frac{1}{(\ln t)^3} \right) \right], \quad as \quad t \to \infty. \tag{6}$$

▸ Go to Theorems Proof

Benktander type II distribution

1. Definition.
This distribution is near to the Gnedenko-Weibull. The formula Benktander type II average residual resource looks like the asymptotics of the average residual resource of the Weibull distribution. But it is an exact equality, yet for the residual variance there is only an asymptotic expansion. Consider the Benktander type II distribution with distribution tail

$$1 - F(t) = e^{\frac{\alpha}{\beta}} t^{\beta-1} e^{-\frac{\alpha}{\beta} t^{\beta}} \quad \alpha > 0,\ 0 < \beta < 1. \tag{7}$$

Benktander type II distribution. Moments

1. It is immediately follows from (1) (see Table 3.4.7 , [**Em**])

$$\mu(t) = \frac{1}{\alpha} t^{1-\beta}, \ \alpha > 0, \ 0 < \beta < 1. \tag{8}$$

2. Residual variance:

$$\sigma^2(t) = \frac{t^{2-2\beta}}{\alpha^2} \left(1 - 2\frac{(\beta - 1)}{\alpha} t^{-\beta} + 2\frac{(\beta - 1)(2\beta - 1)}{\alpha^2} t^{-2\beta} + O\left(t^{-3\beta}\right)\right) \quad t \to \infty, \tag{9}$$

Burr distribution

1. Definition.
The Burr distribution or Burr Type XII distribution ([**Rd**]) is popular as a fit model of a set insurance data:

$$F(t) = 1 - \frac{1}{(1 + t^c)^k}, \quad c > 0, \quad k > 0, \quad t \geq 0.$$

Burr distribution

$$\mu(t) = \frac{t}{ck-1}\left[1 + \frac{ck}{ck-1+c}\cdot\frac{1}{t^c} + \frac{ck(1-c)}{(ck-1+c)(ck-1+2c)}\cdot\frac{1}{t^{2c}} + o\left(\frac{1}{t^{2c}}\right)\right];$$

$$\sigma^2(t) = \frac{ck}{(ck-1)^2}\cdot t^2\cdot\left[\frac{1}{ck-2} + \frac{2c(k+1)}{(ck-1+c)(ck-2+c)}\cdot\frac{1}{t^c} + o\left(\frac{1}{t^c}\right)\right],\quad t\to\infty.$$

Conclusion

We illustrate the behavior of the empirical composite samples for some simulated data sets of heavy-tailed distributions in Section 1. Explicit accurate innovated asymptotic expansions are found of the MRL (ME) and residual variance for well known models. By application of criterion of Balkema; de Haan [**Blk**] one may found the belongs (or not) of these distributions to domain of attraction exponential distribution.

[Blk] Balkema, A.A.; de Haan, L.,*Residual life time at great age. The Annals of probability*, **2** (5), 792–804, (1974).

Thank You for Your Attention

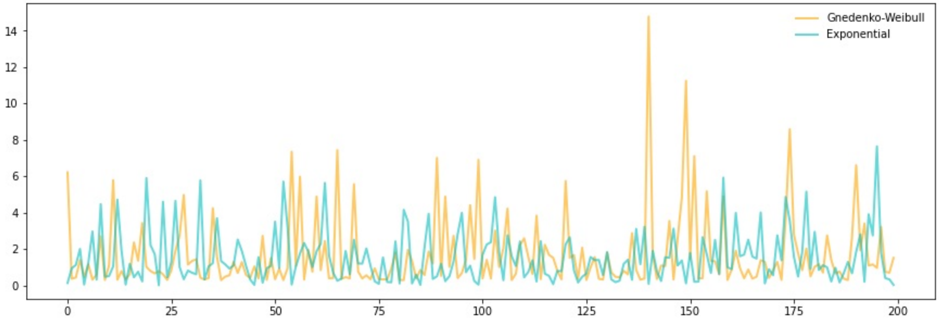
- Embrechts P., Klüppelberg C., Mikosch T. Modelling Extremal Events for Insurance and Finance.; Publisher: Berlin, Springer, 1997.
- Rusev, V.; Skorikov, A. Residual Life Time of the Gnedenko Extreme - Value Distributions, Asymptotic Behavior and Applications. *Shiryaev A.N., Samouylov K.E., Kozyrev D.V. (eds) Recent Developments in Stochastic Methods and Applications. ICSM-5 2020. Springer Proceedings in Mathematics and Statistics* **2021**, *Vol.371*, 292–305.
- Willinger W., Taqqu M. S., Sherman R. and Wilson D.V. Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level. *IEEE/ACM Transactions on Networking*, **1997**, *Vol.5, No. 1*, 71–86.

References II

- Dengaev A.V., Rusev V.N., Skorikov A.V. The Mean Residual Life (MRL) of the Gnedenko-Weibull Distribution. Estimates of Residual Life Time of Pump Submersible Equipment. *Proceedings of Gubkin Russian State University of Oil and Gas* **2020**, *1/298*, 25–37.
- Gradshteyn I. R., Ryzhik I.M. *Tables of Integrals, Series, and Products*; Publisher:Oxford: Elsevier Inc., 2007.
- Rodriges, R. A guide to the Burr type XII distributions. *Biometrika.* **1977**, *Vol. 64, Issue 1, April 1977*, 129–134.
- Balkema, A.A.; de Haan, L. Residual life time at great age. *The Annals of probability* **1974**, *Vol.2, No.5*, 792–804.

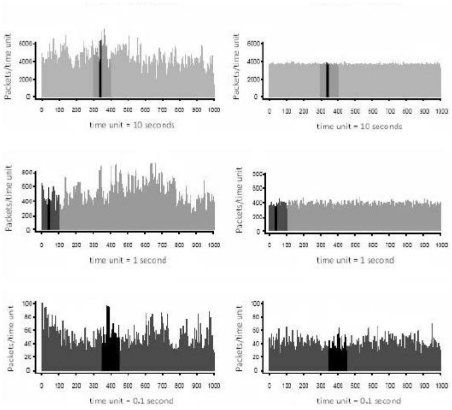
Appendix - A figure

Synthetic plot generated from comparable composite samples of Gnedenko-Weibull distribution with shape parameter 0.69 and scale parameter 1 and exponential distribution with parameter 0,64



Return to presentation

Self-Similar Ethernet Traffic vs Synthetic Traffic generated from a comparable compound Poisson. (Jagerman D.L., Melamed B., Willinger W., 1996)



Return to presentation

Appendix - Terms

Some Abbreviations:

- Cumulative distribution function: DF
- Mean residual life: MRL
- Mean excess function:ME
- Probability density function: PDF

Return to presentation

1. *Proof.* Note that (5) is contained in the Table 3.4.7 , **Embr**, p.161 without proof. The proof is given for fullness.
2. From the above result and the presentation for the integral

$$\int_x^\infty e^{-z^2}dz = e^{-x^2}\left(\frac{1}{2x} + \frac{1}{4x^3} + O\left(\frac{1}{x^5}\right)\right), \quad x \to \infty$$

it follows

$$\int_{t}^{+\infty} (1 - F(x)) \cdot \mu(x)\, dx = \frac{1}{\alpha} \int_{t}^{+\infty} e^{-\beta(\ln x)^2 - (\alpha+1)\ln x} x dx =$$

$$= \frac{1}{\alpha} \int_{\ln t}^{\infty} e^{-\beta(z)^2 - (\alpha+1)z} e^{z} dz = \frac{1}{\alpha} \int_{\ln t}^{\infty} e^{-\beta(z)^2 - \alpha z + z} dz =$$

$$= \frac{e^{\frac{(\alpha-1)^2}{4\beta}}}{\alpha} \int_{\ln t}^{\infty} e^{-\beta\left(z + \frac{\alpha-1}{2\beta}\right)^2} d\left(z + \frac{\alpha-1}{2\beta}\right) = \frac{e^{\frac{(\alpha-1)^2}{4\beta}}}{\alpha\sqrt{\beta}} \int_{\sqrt{\beta}\left(\ln t + \frac{\alpha-1}{2\beta}\right)}^{\infty} e^{-u^2} du =$$

$$= \frac{e^{\frac{(\alpha-1)^2}{4\beta}}}{\alpha\sqrt{\beta}} e^{-\beta\left(\ln t + \frac{\alpha-1}{2\beta}\right)^2} \left[\frac{1}{\sqrt{\beta}\left(\ln t + \frac{\alpha-1}{2\beta}\right)} + \frac{1}{4\left(\sqrt{\beta}\left(\ln t + \frac{\alpha-1}{2\beta}\right)\right)^3} + \cdots \right].$$

Residual variance according as (2) and above asymptotics is derives by equation

$$\sigma^2(t) = \frac{2\alpha}{\alpha + 2\beta \ln t} \cdot e^{\beta(\ln t)^2 + (\alpha+1)\ln t} \cdot \frac{e^{\frac{(\alpha-1)^2}{4\beta}}}{\alpha\sqrt{\beta}} e^{-\beta\left(\ln t + \frac{\alpha-1}{2\beta}\right)^2} \times$$

$$\times \left[\frac{1}{2\sqrt{\beta}\left(\ln t + \frac{\alpha-1}{2\beta}\right)} + \frac{1}{4\left(\sqrt{\beta}\left(\ln t + \frac{\alpha-1}{2\beta}\right)\right)^3} + O\left(\frac{1}{(\ln t)^5}\right) \right] - \frac{t^2}{(\alpha + 2\beta \ln t)^2} =$$

$$= \frac{2t^2}{\alpha + 2\beta \ln t} \cdot \left[\frac{1}{(2\beta \ln t + \alpha - 1)} + \frac{1}{4\beta\left(\ln t + \frac{\alpha-1}{2\beta}\right)^3} + O\left(\frac{1}{(\ln t)^5}\right) \right] - \frac{t^2}{(\alpha + 2\beta \ln t)^2} =$$

Appendix - Proof continue 4

$$= \frac{t^2}{\alpha + 2\beta \ln t}\left[\frac{2}{(2\beta \ln t + \alpha - 1)} - \frac{1}{(2\beta \ln t + \alpha)} + \frac{4\beta^2}{(2\beta \ln t + \alpha - 1)^3} + O\left(\frac{1}{(\ln t)^5}\right)\right] =$$

$$= \frac{t^2}{(\alpha + 2\beta \ln t)^2}\left[\frac{2\beta \ln t + \alpha + 1}{(2\beta \ln t + \alpha - 1)(2\beta \ln t + \alpha)} + O\left(\frac{1}{(\ln t)^3}\right)\right] =$$

$$= \frac{t^2}{(\alpha + 2\beta \ln t)^2}\left[1 + \frac{2}{2\beta \ln t + \alpha - 1} + O\left(\frac{1}{(\ln t)^3}\right)\right].$$

Return to presentation