BULLETIN OF THE INTERNATIONAL MATHEMATICAL VIRTUAL INSTITUTE
ISSN (p) 2303-4874, ISSN (o) 2303-4955
www.imvibl.org /JOURNALS / BULLETIN
Vol. 6(2016), 227-234

COVERAGE PROBABILITY FOR ONE QUANTILE ESTIMATOR AND ITS APPLICATION ON INSURANCE DATA

Jelena Stanojević and Marija Jovović

Communicated by Branislav Boričić

Abstract. Rate of coverage probability for direct-simulation estimator of quantile estimation has been determined for Pareto 1, Pareto 2 and general case, for the general Pareto distribution. Appropriate calculations and tables with results have been given for data with Pareto 2 distribution as well as for real data from insurance analyses which can be fit with that distribution.

1. Introduction

Finding high quantile estimators of an unknown distribution function is a problem of great importance from both a theoretical as well as a practical perspective. For example, important problem that involves high quantile estimators is the estimation of the Value-at-Risk (VaR), which is a standard risk measure in the field of finance. Many authors have given different estimators of high quantiles and usually they have used k upper order statistics of a sample, see [**5**], [**9**], [**4**], [**6**] and references therein. The problem of appropriate choice of k is complicated, because small k gives the estimator with large deviation and large k gives the estimator with large bias, and it is still open problem for another paper. In this paper we consider well known quantile estimator from the literature, the direct-simulation estimator of large quantiles and its rate of convergence. One of the aims of the paper is to calculate that rate of convergence of the Pareto distribution, considering its multiple applications in empirical analyzes and in theory. For example, certain socio-economic quantities, magnitudes of earthquakes, number of hits at websites, the assets of firms as well as standardized price returns on individual stocks or stock indices are described with the Pareto distribution. For references see [**3**], [**12**], [**16**],

2010 *Mathematics Subject Classification.* 62G05, 62G30, 62G99, 62P05.

Key words and phrases. high quantile estimation, VaR, direct-simulation estimator, coverage probability.

This paper is partially supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia No.174020 and No.179005.

[**8**], [**13**], [**1**], [**7**]. Also, the Pareto distribution is widely used in various fields of science, such as hydrology, geology, climatology, astronomy, physics, finance and for this paper the most important is insurance (see [**14**], [**2**], [**10**]).
In many types of insurance, loss data contain observations with high intensity and low frequency. Since the appropriate statistical distributions for modeling insurance data are skewed, high quantile estimators of these distributions could be used as the adequate measures of actuarial risks. Value-at-Risk at a sufficiently high confidence level determines probable maximum loss as well as solvency capital requirements of insurance companies. The Solvency II framework adopted Value-at-Risk at a 99.5% confidence level over a one year period as a measure of all risks threatening financial health of insurers operating in the European Union (see [**15**], [**11**]). In the light of the new methodological approach to evaluating insurance companies' solvency, the issue of this quantile estimation gains growing importance in the insurance sector nowadays.

2. Some Preliminaries and Notations

Let $X_1, ..., X_n$ are i.i.d. random variables with common nondegenerate distribution function $F(x)$. The empirical distribution function is defined with $F_n(x) = \frac{1}{n}\sum_{k=1}^{n} I(X_k \leqslant x)$, $x \in R$, where $I(X_k \leqslant x)$ denotes the indicator of the event $\{X_k \leqslant x\}$. x_p is appropriate quantile of the distribution function $F(x)$ and it is defined with: $x_p = \inf\{x : F(x) \geqslant p\}$, $p \in (0,1)$. In this paper we will consider the direct-simulation estimator, defined with formula: $\widehat{x}_p(n) = \inf\{t : F_n(t) \geqslant p\}$, already known in the literature. The term of negative dependence is important for understanding of the next section. Random variables $X_1, X_2, \ldots, X_n$ are negatively dependent if the following two inequalities hold for all $x_1, x_2, ..., x_n$: $P\{X_1 \leqslant x_1, ..., X_n \leqslant x_n\} \leqslant P\{X_1 \leqslant x_1\} \cdot ... \cdot P\{X_n \leqslant x_n\}$, $P\{X_1 \geqslant x_1, ..., X_n \geqslant x_n\} \leqslant P\{X_1 \geqslant x_1\} \cdot ... \cdot P\{X_n \geqslant x_n\}$. Two lemmas which are also important for understanding of the next section can be found in [**18**]. Here we will give only the primary theorem from that paper.

Theorem 2.1. ([**18**]) *If the distribution function F is strictly increasing and $\{Y_n, n \geqslant 1\}$ are negatively dependent, then*

$$P\{|\widehat{x}_p(n) - x_p| \geqslant \epsilon\} \leqslant e^{-n\Delta_+(\epsilon,n)} + e^{-n\Delta_-(\epsilon,n)}, \quad \textit{for all } \epsilon > 0, \tag{2.1}$$

where

$$\Delta_+(\epsilon, n) = \sup_{-\infty<\lambda\leqslant 0}\Big(\lambda p - \frac{\ln E\{\exp[\lambda \sum_{i=1}^{n} I(Y_i \leqslant x_p + \epsilon)]\}}{n}\Big),$$

$$\Delta_-(\epsilon, n) = \sup_{0\leqslant\lambda<+\infty}\Big(\lambda p - \frac{\ln E\{\exp[\lambda \sum_{i=1}^{n} I(Y_i \leqslant x_p - \epsilon)]\}}{n}\Big).$$

And, moreover, the rate is enhanced by negatively dependence in the sense that

$$\Delta_+(\epsilon, n) \geqslant \sup_{-\infty<\lambda\leqslant 0}(\lambda p - \ln E\{\exp[\lambda I(Y \leqslant x_p + \epsilon)]\}) > 0,$$

$$\Delta_-(\epsilon, n) \geqslant \sup_{0\leqslant\lambda<+\infty}(\lambda p - \ln E\{\exp[\lambda I(Y \leqslant x_p - \epsilon)]\}) > 0,$$

where the right-hand "sup" are the rates for i.i.d. samples.

3. Results for Different Types of Pareto Distribution

In this section we will determine the rate of convergence of the direct-simulation estimator $\widehat{x}_p(n)$ of the quantile x_p to exact value, for different types of Pareto distribution.
For Pareto 1 distribution function the density function is $f(x) = ab^a x^{-(a+1)}$ and the distribution function is $F(x) = 1 - b^a x^{-a}$, with $\overline{F}(x) = b^a x^{-a}$, $x \geqslant b$, for shape parameter $a > 0$ and scale parameter $b > 0$. The next result is given in [**17**].

Theorem 3.1. ([**17**]) *Let $\{Y_n, n \geqslant 1\}$ be negatively dependent random variables with the common Pareto 1 distribution $\overline{F}(x) = b^a x^{-a}$, $x \geqslant b$, $a, b > 0$. The rate of convergence for the standard quantile estimator $\widehat{x}_p(n)$ in this case is given by*

$$\begin{aligned}(3.1)\qquad P\ &\{\left|\widehat{x}_p(n) - x_p\right| \geqslant \epsilon\} \leqslant e^{-n\Delta_+} + e^{-n\Delta_-}\\ =\ &\Big(\frac{pb^a(x_p+\epsilon)^{-a}}{(1-b^a(x_p+\epsilon)^{-a})(1-p)}\Big)^{-np} \cdot \Big(\frac{b^a(x_p+\epsilon)^{-a}}{1-p}\Big)^n\\ +\ &\Big(\frac{pb^a(x_p-\epsilon)^{-a}}{(1-b^a(x_p-\epsilon)^{-a})(1-p)}\Big)^{-np} \cdot \Big(\frac{b^a(x_p-\epsilon)^{-a}}{1-p}\Big)^n,\end{aligned}$$

and we may write:

$$\begin{aligned}P\{\left|\widehat{x}_p(n) - x_p\right| \geqslant \epsilon\} \ &\leqslant\ \Big(\frac{p(1-p_+)}{p_+(1-p)}\Big)^{-np} \cdot \Big(\frac{1-p_+}{1-p}\Big)^n\\ &+\ \Big(\frac{p(1-p_-)}{p_-(1-p)}\Big)^{-np} \cdot \Big(\frac{1-p_-}{1-p}\Big)^n,\end{aligned}$$

for $1 - p_+ = b^a(x_p+\epsilon)^{-a}$ and $1 - p_- = b^a(x_p-\epsilon)^{-a}$.

Another special case is Pareto 2 distribution, with the density function $f(x) = ab^a(x+b)^{-(a+1)}$ and the distribution function $F(x) = 1 - b^a(x+b)^{-a}$, with $\overline{F}(x) = b^a(x+b)^{-a}$, $x \geqslant 0$, for shape parameter $a > 0$ and scale parameter $b > 0$. In this case we have the next result.

Theorem 3.2. *Let $\{Y_n, n \geqslant 1\}$ be negatively dependent random variables with the common Pareto 2 distribution, $\overline{F}(x) = b^a(x+b)^{-a}$, $x \geqslant 0$, $a, b > 0$. The rate of convergence for the standard quantile estimator $\widehat{x}_p(n)$ in this case is given by*

$$\begin{aligned}(3.2)\qquad P\ &\{\left|\widehat{x}_p(n) - x_p\right| \geqslant \epsilon\} \leqslant e^{-n\Delta_+} + e^{-n\Delta_-}\\ =\ &\Big(\frac{pb^a(x_p+\epsilon+b)^{-a}}{(1-b^a(x_p+\epsilon+b)^{-a})(1-p)}\Big)^{-np} \cdot \Big(\frac{b^a(x_p+\epsilon+b)^{-a}}{1-p}\Big)^n\\ +\ &\Big(\frac{pb^a(x_p-\epsilon+b)^{-a}}{(1-b^a(x_p-\epsilon+b)^{-a})(1-p)}\Big)^{-np} \cdot \Big(\frac{b^a(x_p-\epsilon+b)^{-a}}{1-p}\Big)^n,\end{aligned}$$

and we may write:

$$\begin{aligned} P\{\left|\widehat{x}_p(n)-x_p\right| \geqslant \epsilon\} &\leqslant \Big(\frac{p(1-p_+)}{p_+(1-p)}\Big)^{-np}\cdot\Big(\frac{1-p_+}{1-p}\Big)^{n} \\ &+ \Big(\frac{p(1-p_-)}{p_-(1-p)}\Big)^{-np}\cdot\Big(\frac{1-p_-}{1-p}\Big)^{n}, \end{aligned}$$

for $1-p_+ = b^a(x_p+\epsilon+b)^{-a}$ *and* $1-p_- = b^a(x_p-\epsilon+b)^{-a}$.

Proof. Since the Pareto 2 distribution is strictly increasing we may use Theorem (2.1) and obtain:

$$P\{\left|\widehat{x}_p(n)-x_p\right| \geqslant \epsilon\} \leqslant e^{-n\Delta_+(\epsilon,n)}+e^{-n\Delta_-(\epsilon,n)}, \quad \text{for all } \epsilon>0,$$

where

$$\begin{aligned} \Delta_+(\epsilon,n) &\geqslant \sup_{-\infty<\lambda\leqslant 0}(\lambda p-\ln E\{\exp[\lambda I(Y\leqslant x_p+\epsilon)]\}) = \Delta_+, \\ \Delta_-(\epsilon,n) &\geqslant \sup_{0\leqslant\lambda<+\infty}(\lambda p-\ln E\{\exp[\lambda I(Y\leqslant x_p-\epsilon)]\}) = \Delta_-, \\ p &= P[Y\leqslant x_p]. \end{aligned}$$

If we denote, $p_+ = P[Y\leqslant x_p+\epsilon] = F(x_p+\epsilon) = 1-b^a(x_p+\epsilon+b)^{-a}$ and $p_- = P[Y\leqslant x_p-\epsilon] = F(x_p-\epsilon) = 1-b^a(x_p-\epsilon+b)^{-a}$, our goal is to obtain Δ_+ and Δ_-.

The distribution of the indicator $I(Y\leqslant x_p+\epsilon)$ is given by

$$I(Y\leqslant x_p+\epsilon):\left(\begin{array}{cc} 0 & 1 \\ 1-p_+ & p_+ \end{array}\right).$$

Now, we may calculate

$$\Delta_+ = \sup_{-\infty<\lambda\leqslant 0}(\lambda p-\ln(e^{\lambda}p_+ + 1-p_+)).$$

The maximum of the above function is attained for $\lambda = \ln\frac{p(1-p_+)}{p_+(1-p)}$ and λ is always negative (since $p<p_+$). Consequently, we obtaine

$$\Delta_+ = p\ln\frac{p(1-p_+)}{p_+(1-p)} - \ln\frac{1-p_+}{1-p}.$$

Similarly, we may calculate

$$\Delta_- = \sup_{0\leqslant\lambda<+\infty}(\lambda p-\ln(e^{\lambda}p_- + 1-p_-)).$$

The maximum of the above function is attained for $\lambda = \ln\frac{p(1-p_-)}{(1-p)p_-}$ and λ is always positive (since $p>p_-$). Consequently, we obtain

$$\Delta_- = p\ln\frac{p(1-p_-)}{p_-(1-p)} - \ln\frac{1-p_-}{1-p}.$$

Finally we have the result:

$$\begin{aligned} P\{\left|\widehat{x}_p(n) - x_p\right| \geqslant \epsilon\} &\leqslant e^{-n\Delta_+} + e^{-n\Delta_-} \\ &= \Big(\frac{p(1-p_+)}{p_+(1-p)}\Big)^{-np} \cdot \Big(\frac{1-p_+}{1-p}\Big)^{n} \\ &+ \Big(\frac{p(1-p_-)}{p_-(1-p)}\Big)^{-np} \cdot \Big(\frac{1-p_-}{1-p}\Big)^{n}, \end{aligned}$$

for $1 - p_+ = b^a(x_p + \epsilon + b)^{-a}$ and $1 - p_- = b^a(x_p - \epsilon + b)^{-a}$ and the proof is completed. □

Also we may analyze more general case, for example the general Pareto distribution, $\overline{F}(x) = L(x)x^{-a}$, where $a > 0$ and $L(x)$ is slowly varying function, meaning:

$$\lim_{t\to+\infty} \frac{L(tx)}{L(t)} = 1.$$

In this case we can obtain the next result:

$$\begin{aligned} \text{(3.3)}\quad P\quad &\{\left|\widehat{x}_p(n) - x_p\right| \geqslant \epsilon\} \leqslant e^{-n\Delta_+} + e^{-n\Delta_-} \\ = \quad &\Big(\frac{pL(x_p+\epsilon)(x_p+\epsilon)^{-a}}{(1-p)(1-L(x_p+\epsilon)(x_p+\epsilon)^{-a})}\Big)^{-np} \cdot \Big(\frac{L(x_p+\epsilon)(x_p+\epsilon)^{-a}}{1-p}\Big)^{n} \\ + \quad &\Big(\frac{pL(x_p-\epsilon)(x_p-\epsilon)^{-\alpha}}{(1-p)(1-L(x_p-\epsilon)(x_p-\epsilon)^{-\alpha})}\Big)^{-np} \cdot \Big(\frac{L(x_p-\epsilon)(x_p-\epsilon)^{-a}}{1-p}\Big)^{n}, \end{aligned}$$

which is analogous with the well known result:

$$\begin{aligned} P\{\left|\widehat{x}_p(n) - x_p\right| \geqslant \epsilon\} &\leqslant \Big(\frac{p(1-p_+)}{p_+(1-p)}\Big)^{-np} \cdot \Big(\frac{1-p_+}{1-p}\Big)^{n} \\ &+ \Big(\frac{p(1-p_-)}{p_-(1-p)}\Big)^{-np} \cdot \Big(\frac{1-p_-}{1-p}\Big)^{n}, \end{aligned}$$

for $1 - p_+ = L(x_p + \epsilon)(x_p + \epsilon)^{-a}$ and $1 - p_- = L(x_p - \epsilon)(x_p - \epsilon)^{-a}$. The proof in this case is analogous as the proof for the Pareto 2 distribution and we will omit it here.

4. Numerical Examples and Real Data Analysis

In this section we present some numerical examples and performance of formulas from the section above. We take two values of b and a and two values of ϵ for the Pareto 2 distribution. For each parameter setting we compute rate of convergence of the direct-simulation estimator $\widehat{x}_p(n)$ of the quantile x_p to exact quantile value, by using Theorem(3.2) and appropriate formula (3.2). Tables 4.1-4.2 contain this results.

Table 4.1 Rate of convergence for Pareto 2 distribution and $a = 4$, $b = 5$, $p = 0.05$, $x_p = 0.06453$

	n	$e^{-n\Delta_+} + e^{-n\Delta_-}$
$\epsilon = 0.05$	50	0.77925
	100	0.40135
	200	0.13988
	300	0.05189
	500	0.00721
$\epsilon = 0.005$	5000	0.95208
	9000	0.52633
	20000	0.10363
	25000	0.04963
	40000	0.0055

Table 4.2 Rate of convergence for Pareto 2 distribution and $a = 6$, $b = 5$, $p = 0.95$, $x_p = 3.237745$

	n	$e^{-n\Delta_+} + e^{-n\Delta_-}$
$\epsilon = 0.5$	300	0.694432
	500	0.351129
	800	0.130302
	1000	0.068538
	1800	0.005822
$\epsilon = 1$	50	0.94848
	100	0.487529
	200	0.152777
	300	0.053786
	500	0.007338

The results in the tables above show the convergence rate of the coverage probability as a function of sample size (and other parameters). It is clear that it decreases as sample size increases. We could see from tables that for small quantile $x_{0.05}$ and for large quantile $x_{0.95}$ the interval for sample size is from 50 to 500, what gives the appropriate results in both cases.

Important question in this moment is: How the proposed results can be used in practice? To answer on that question we analyzed real data set in this section. Our real data sample contains 652 observations for loss amounts (in million RSD) recorded in property insurance in the portfolio of one insurance company operating in Serbia during 2014. Since the Kolmogorov-Smirnov statistic (0.03469) is smaller than the critical value (0.0479) (for $\alpha = 0.1$ and for smaller α is the same result), we cannot reject the hypothesis that this sample stems from the Pareto 2 distribution. The estimated values of distribution parameters are $\widehat{a} = 5.7401$ and $\widehat{b} = 5.0333$. For $p = 0.95$ and $n = 652$ we obtained that $\widehat{x}_{0.95}(652) = 3.5857$. It is possible to calculate quantile x_p for Pareto 2 distribution with parameters $a = 5.7401$ and

$b = 5.0333$ and probability $p = 0.95$ and it is $x_{0.95} = 3.448908$. We obtained that there is less than 28.39% chance that quantile $x_{0.95}$ differs from the direct-simulation estimator $\widehat{x}_{0.95}(652)$ for more than 0.5 and also there is less than 2.25% chance that this difference is greater than 0.8. Two above results we are obtained by using inequality (3.2).

5. Conclusion

In this paper we considered appropriate coverage probability of the direct-simulation estimator $\widehat{x}_p(n)$ of a large quantile x_p and we gave some results for Pareto 1, Pareto 2 and the general Pareto distributions. We take two values of parameters for Pareto 2 distribution (b and a) and two values of ϵ and for each parameter setting we computed rate of convergence of the direct-simulation estimator of the quantile to exact value of quantile, by using appropriate theorem and formula. Tables 4.1-4.2 contain this results and show that rate of convergence decreases as sample increases, what we could expect at the beginning. In the last section, we analyzed real data set and performed appropriate calculations by using main results of this paper. The obtained results could be used to enhance determination of the maximum probable loss and solvency capital requirements of insurance companies.

Acknowledgement

The authors would like to thank the reviewer for useful remarks which have improved this paper.

References

[1] Aoyama, H., Souma, W. and Fujiwara Y., Growth and fluctuations of personal and company's income. *Physica A: Statistical Mechanics and its Applications*, **324** (1-2) (2003), 352-358.

[2] Burnecki, K., Kukla, G. and Weron, R., Property insurance loss distributions. *Physica A: Statistical Mechanics and its Applications*, **287** (1-2) (2000), 269-278.

[3] Champernowne, D., A model of income distribution. *Economic Journal*, **63**(250) (1953), 318-351.

[4] Dekkers, A.R.M and de Haan, L., On the estimation of the extreme value index and large quantile estimation. *Annals of Statistics*, **17**(4) (1989), 1795-1832.

[5] Embrechts, P., Kluppelberg, C. and Mikosch, T., *Modelling Extremal Events for Insurance and Finance*. Springer 1997.

[6] Feldman, D. and Tucker, H.G., Estimation of non-unique quantiles. *Annals of Mathematical Statistics*, **37**(2) (1966), 451-457.

[7] Fujiwara, Y., Aoyama, H., Di Guilmi, C., Souma, W. and Gallegati, M., Gibrat and Pareto-Zipf revisited with European firms. *Physica A: Statistical Mechanics and its Applications*, **344**(1-2) (2004), 112-116.

[8] Levy, M. and Solomon, S., New evidence for the power-law distribution of wealth. *Physica A: Statistical Mechanics and its Applications*, **242**(1-2) (1997), 90-94.

[9] Matthys, G. and Beirlant, J., Estimating the extreme value index and high quantiles with exponential regression models. *Statistica Sinica*, **13** (2003), 853-880.

[10] Matthys, G., Delafosse, E., Guillou, A. and Beirlant, J., Estimating catastrophic quantile levels for heavy-tailed distributions. *Insurance: Mathematics and Economics*, **34**(3) (2004), 517-537.

[11] Pfeifer, D. and Strassburger, D., Solvency II: stability problems with the SCR aggregation formula. *Scandinavian Actuarial Journal*, **2008**(1) (2008), 61-77.

[12] Quandt, R.E., Old and New Methods of Estimation and the Pareto Distribution. *Metrika*, Vol. **10**, No.1 (1966), 55-82.

[13] Reed, W. J., The Pareto, Zipf and other power laws. *Economics Letters*, **74**(1) (2001), 15-19.

[14] Resnick, S., Discussion of the Danish Data on Large Fire Insurance Losses. *ASTIN Bulletin*, **27**(1) (1997), 139-151.

[15] Sandström, A., Solvency II: Calibration for skewness. *Scandinavian Actuarial Journal*, **2007**(2) (2007), 126-134.

[16] Singh, S.K. and Maddala, G.S., A Function for Size Distribution of Incomes. *Econometrica*, **44**(5) (1976), 963-970.

[17] Stanojević, J., On estimation of high quantiles for certain classes of distributions. *Yugoslav Journal of Operations Research*, Vol. **25**(2) (2014), 299-312.

[18] Jin, X. and Fu, M.C., *A Large Deviations Analysis of Quantile Estimation with Application to Value at Risk*, MIT Operations Research Center Seminar Series (2002).

Received by editors 22.02.2015; Revised version 20.02.2016. Available on line 29.02.2016.

Department of Mathematics and Statistics, Faculty of Economics, University of Belgrade, Serbia
E-mail address: jelenas@ekof.bg.ac.rs

Department of Economic Policy and Development, Faculty of Economics, University of Belgrade, Serbia
E-mail address: marijajovovic@ekof.bg.ac.rs