

## THEORETICAL AND EMPIRICAL RESULTS ON THE COVERAGE PROBABILITY OF THE STANDARD QUANTILE ESTIMATOR FOR THE BURR DISTRIBUTION

JELENA STANOJEVIĆ<sup>1</sup>, MARIJA KOPRIVICA<sup>2</sup>

<sup>1</sup> University of Belgrade, Faculty of Economics and Business, Belgrade, [jelena.stanojevic@ekof.bg.ac.rs](mailto:jelena.stanojevic@ekof.bg.ac.rs),  0000-0001-5668-5297

<sup>2</sup> University of Belgrade, Faculty of Economics and Business, Belgrade, [marija.koprivica@ekof.bg.ac.rs](mailto:marija.koprivica@ekof.bg.ac.rs),  0000-0003-4239-2252

**Abstract:** *This In this paper, we determine the coverage probability of the standard quantile estimator for the Burr Type XII distribution. The analysis is conducted under the assumptions of negative dependence and independence and identical distribution (i.i.d.) between observations, and is based on large deviations theory. Numerical results are provided for various parameter settings and quantile levels, illustrating the estimator's convergence properties as the sample size increases. Our findings improve the understanding of the estimator's performance in the context of heavy-tailed distributions relevant to insurance and finance.*

**Keywords:** Coverage probability, standard quantile estimator, Value-at-Risk

### 1. INTRODUCTION

The estimation of high quantiles plays a crucial role in many fields, particularly in risk management, finance, and insurance. High quantiles — such as the 95th or 99.5th percentiles — are especially important in extreme value analysis and are commonly used to define risk measures, such as the widely accepted Value-at-Risk (VaR). Accurate quantile estimation is essential for developing reliable statistical models to evaluate extreme events.

Most of the literature on quantile estimation focuses on symmetric or moderately skewed distributions (Miljković, Causey, and Jovanović, 2022). However, in practice, insurance and financial loss data often exhibit extreme skewness and heavy tails. This is particularly relevant for lines of insurance prone to catastrophic claims, such as property, aviation, and marine insurance. The Burr Type XII distribution — a three-parameter, heavy-tailed distribution (Burr, 1942) — has gained attention as a suitable model for such loss data (Beirlant et al., 2001; Cummins et al., 1990; Miljković and Grün, 2016; Venter, 1983). Despite its practical relevance, theoretical results on high quantile estimation under the Burr Type XII distribution, especially in the presence of dependence structures, remain limited.

The main goal of the paper is to derive the upper bound for the probability  $P\{|\hat{x}_p(n) - x_p| \geq \varepsilon\}$ , which we will denote as  $\eta$ . However, the probability of the true quantile value  $x_p$  belonging to the random interval  $|\hat{x}_p(n) - \varepsilon, \hat{x}_p(n) + \varepsilon|$ , i.e. the "coverage probability" is equal to  $1 - \eta$ . The quantity  $\eta$  can be understood as the convergence rate (in probability) for any fixed  $\varepsilon > 0$ . We analyze the upper bound for the convergence rate, or, equivalently, the lower bound for the coverage probability, for a given  $\varepsilon$ , of the standard quantile estimator for the Burr Type XII distribution. The analysis is based on large deviations theory (Jin and Fu, 2007), and is conducted under the assumptions of independence and identical distribution (i.i.d.) between observations, however the results may be used in the case of negatively dependent variables as well. We also complement the theoretical results with numerical results to illustrate the impact of parameter settings and sample size on convergence.

### 2. SOME DEFINITIONS AND PRELIMINARIES

Let  $(X_1, X_2, \dots, X_n)$  be a sample from a random variable  $X$  with a nondegenerate distribution  $F(x)$ . The empirical distribution function of  $X$  is given by:

$$F_n(x) = \frac{1}{n} \sum_{k=1}^n I\{X_k \leq x\}, \quad x \in R,$$

where  $I\{X_k \leq x\}$  denotes the indicator function of the event  $\{X_k \leq x\}$ .

➤ **Definition 1.** The  $p$ -th quantile  $x_p$  of the distribution  $F(x)$  is defined as:

$$x_p = \inf\{x: F(x) \geq p\}, \quad p \in (0,1). \quad (1)$$

In this paper, we consider the case of the direct-simulation quantile estimator, which is the standard quantile estimator in the literature, and obtain the main result for this setting.

➤ **Definition 2.** The direct-simulation quantile estimator of  $X$  is defined by the formula:

$$\hat{x}_p(n) = \inf\{t: F_n(t) \geq p\}, \quad p \in (0,1), \quad (2)$$

where  $F_n(\cdot)$  is the empirical distribution function of  $X$ .

Another important concept in this paper is negative dependence.

➤ **Definition 3.** Random variables  $X_1, X_2, \dots, X_n$  are said to be negatively dependent if the following two inequalities hold for all  $x_1, x_2, \dots, x_n$ :

$$P\{X_1 \leq x_1, \dots, X_n \leq x_n\} \leq P\{X_1 \leq x_1\} \cdots P\{X_n \leq x_n\},$$

and

$$P\{X_1 \geq x_1, \dots, X_n \geq x_n\} \leq P\{X_1 \geq x_1\} \cdots P\{X_n \geq x_n\}.$$

An auxiliary result used in this paper is the following theorem, which is the main result in (Jin and Fu, 2007)

➤ **Theorem 4.** If the distribution  $F$  is strictly increasing and the sequence  $\{X_n, n \geq 1\}$  consists of negatively dependent random variables, then

$$P\{|\hat{x}_p(n) - x_p| \geq \varepsilon\} \leq e^{-n\Delta_+(\varepsilon, n)} + e^{-n\Delta_-(\varepsilon, n)}, \quad \text{for all } \varepsilon > 0, \quad (3)$$

where

$$\Delta_+(\varepsilon, n) = \sup_{-\infty < \lambda \leq 0} \left( \lambda p - \frac{\ln E \left( \exp(\lambda \sum_{i=1}^n I\{X_i \leq x_p + \varepsilon\}) \right)}{n} \right),$$

$$\Delta_-(\varepsilon, n) = \sup_{0 \leq \lambda < +\infty} \left( \lambda p - \frac{\ln E \left( \exp(\lambda \sum_{i=1}^n I\{X_i \leq x_p - \varepsilon\}) \right)}{n} \right).$$

Moreover, the rate is improved by negative dependence in the sense that

$$\Delta_+(\varepsilon, n) \geq \sup_{-\infty < \lambda \leq 0} \left( \lambda p - \ln E \left( \exp(\lambda I\{X \leq x_p + \varepsilon\}) \right) \right) > 0,$$

$$\Delta_-(\varepsilon, n) \geq \sup_{0 \leq \lambda < +\infty} \left( \lambda p - \ln E \left( \exp(\lambda I\{X \leq x_p - \varepsilon\}) \right) \right) > 0,$$

where the "sup" on the right-hand side denotes the appropriate supremum and represents the rate for i.i.d. samples.

### 3. THE MAIN RESULT

In this section, we investigate a specific case of the Burr distribution and calculate the upper bound for the convergence rate of the standard quantile estimator  $\hat{x}_p(n)$  to the true quantile value  $x_p$ , or, equivalently, the lower bound for the appropriate coverage probability.

Similar results for various forms of the Pareto distribution, including the Generalized Pareto Distribution (GPD), were presented in.

We consider the Burr Type XII distribution, also known as the Singh–Maddala distribution (sometimes referred to as the generalized log-logistic distribution), with the distribution function given by:

$$F(x) = 1 - \left( \frac{\theta}{\theta + x^\tau} \right)^\alpha, \quad x \geq 0,$$

with appropriate parameters  $\alpha, \theta, \tau > 0$ . We obtain the following result.

Let  $\{X_n, n \geq 1\}$  be i.i.d. random variables with a common Burr Type XII distribution, given by  $\bar{F}(x) = \left( \frac{\theta}{\theta + x^\tau} \right)^\alpha = \theta^\alpha (\theta + x^\tau)^{-\alpha}$ ,  $x \geq 0$ , where  $\alpha, \theta, \tau > 0$ . The upper bound for the convergence rate of the standard quantile estimator  $\hat{x}_p(n)$  to the true quantile value  $x_p$ , is given by the following inequality:

$$P\{|\hat{x}_p(n) - x_p| \geq \varepsilon\} \leq \left(\frac{p(1-p_+)}{p_+(1-p)}\right)^{-np} \cdot \left(\frac{1-p_+}{1-p}\right)^n + \left(\frac{p(1-p_-)}{p_-(1-p)}\right)^{-np} \cdot \left(\frac{1-p_-}{1-p}\right)^n, \quad (4)$$

where  $p_+$  and  $p_-$  are:  $p_+ = 1 - \theta^\alpha(\theta + (x_p + \varepsilon)^\tau)^{-\alpha}$  and  $p_- = 1 - \theta^\alpha(\theta + (x_p - \varepsilon)^\tau)^{-\alpha}$ .

*Proof.* The Burr Type XII distribution is strictly increasing, and in the case where  $\{X_n, n \geq 1\}$  are negatively dependent, we can apply Theorem [brzina con]:

$$P\{|\hat{x}_p(n) - x_p| \geq \varepsilon\} \leq e^{-n\Delta_+(\varepsilon, n)} + e^{-n\Delta_-(\varepsilon, n)}, \quad \text{for all } \varepsilon > 0,$$

where

$$\Delta_+(\varepsilon, n) = \sup_{-\infty < \lambda \leq 0} \left( \lambda p - \frac{\ln E \left( \exp(\lambda \sum_{i=1}^n I\{X_i \leq x_p + \varepsilon\}) \right)}{n} \right),$$

$$\Delta_-(\varepsilon, n) = \sup_{0 \leq \lambda < +\infty} \left( \lambda p - \frac{\ln E \left( \exp(\lambda \sum_{i=1}^n I\{X_i \leq x_p - \varepsilon\}) \right)}{n} \right).$$

In the case where  $\{X_n, n \geq 1\}$  are i.i.d., from the same theorem we also have the following result:

$$\Delta_+(\varepsilon, n) \geq \sup_{-\infty < \lambda \leq 0} \left( \lambda p - \ln E \left( \exp(\lambda I\{X \leq x_p + \varepsilon\}) \right) \right),$$

$$\Delta_-(\varepsilon, n) \geq \sup_{0 \leq \lambda < +\infty} \left( \lambda p - \ln E \left( \exp(\lambda I\{X \leq x_p - \varepsilon\}) \right) \right).$$

Let us denote the right-hand sides of the previous inequalities by  $\Delta_+$  and  $\Delta_-$ , respectively, and it is:

$$\Delta_+ = \sup_{-\infty < \lambda \leq 0} \left( \lambda p - \ln E \left( \exp(\lambda I\{X \leq x_p + \varepsilon\}) \right) \right),$$

$$\Delta_- = \sup_{0 \leq \lambda < +\infty} \left( \lambda p - \ln E \left( \exp(\lambda I\{X \leq x_p - \varepsilon\}) \right) \right).$$

If we denote:

$$p_+ = P\{X \leq x_p + \varepsilon\} = F(x_p + \varepsilon) = 1 - \theta^\alpha(\theta + (x_p + \varepsilon)^\tau)^{-\alpha},$$

$$p_- = P\{X \leq x_p - \varepsilon\} = F(x_p - \varepsilon) = 1 - \theta^\alpha(\theta + (x_p - \varepsilon)^\tau)^{-\alpha},$$

$$p = P\{X \leq x_p\},$$

then our goal is to derive expressions for  $\Delta_+$  and  $\Delta_-$ .

The distribution of the indicator  $I\{X \leq x_p + \varepsilon\}$  is given by

$$I\{X \leq x_p + \varepsilon\}: \begin{pmatrix} 0 & 1 \\ 1 - p_+ & p_+ \end{pmatrix},$$

and we can compute:

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

The supremum of the above function is attained at  $\lambda = \ln \frac{p(1-p_+)}{p_+(1-p)}$  (note,  $\lambda$  is always negative since  $p < p_+$ ).

Consequently, we have:

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

Similarly, the distribution of the indicator  $I\{X \leq x_p - \varepsilon\}$  is given by

$$I\{X \leq x_p - \varepsilon\}: \begin{pmatrix} 0 & 1 \\ 1 - p_- & p_- \end{pmatrix},$$

and we can compute:

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

The supremum of the above function is attained at  $\lambda = \ln \frac{p(1-p_-)}{p_-(1-p)}$  (note,  $\lambda$  is always positive since  $p > p_-$ ).

Consequently, we have:

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

Finally,

$$\begin{aligned} P\{|\hat{x}_p(n) - x_p| \geq \varepsilon\} &\leq e^{-n\Delta_+} + e^{-n\Delta_-} \\ &= \left(\frac{p(1-p_+)}{p_+(1-p)}\right)^{-np} \cdot \left(\frac{1-p_+}{1-p}\right)^n + \left(\frac{p(1-p_-)}{p_-(1-p)}\right)^{-np} \cdot \left(\frac{1-p_-}{1-p}\right)^n, \end{aligned}$$

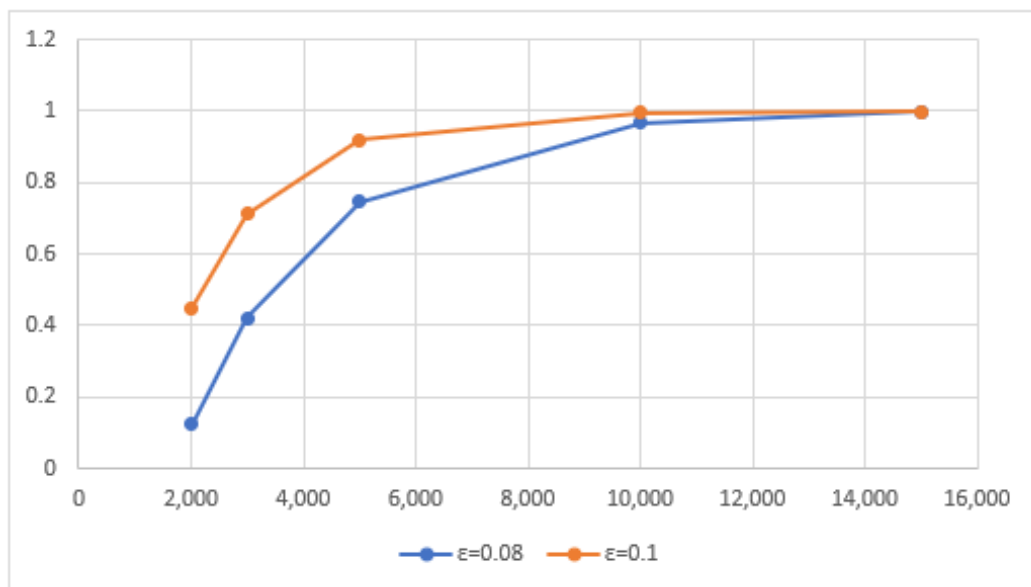
where  $p_+ = 1 - \theta^\alpha(\theta + (x_p + \varepsilon)^\tau)^{-\alpha}$  and  $p_- = 1 - \theta^\alpha(\theta + (x_p - \varepsilon)^\tau)^{-\alpha}$ , and the proof is completed.  $\square$   
 Note, the proof of the above theorem is conducted using the result of upper bound for the convergence rate for the i.i.d. case. Although the bound still holds for the negatively dependent variables, it is looser than the bounds  $\Delta_+(\varepsilon, n)$  and  $\Delta_-(\varepsilon, n)$ , and those refinements are not utilized for the case which we consider here.

#### 4. APPLICATION OF RESULTS

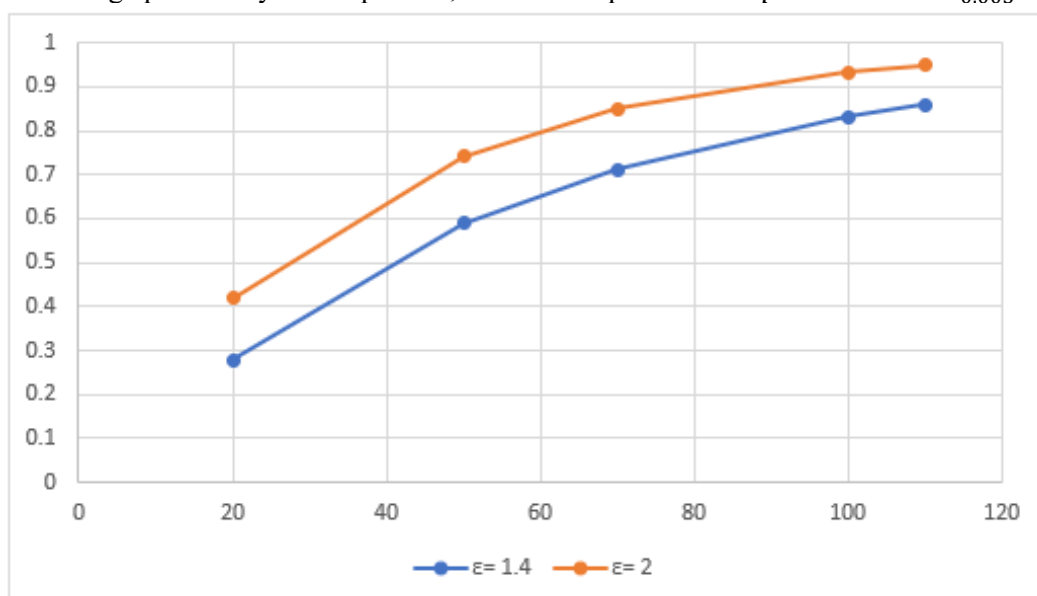
In this section, we present an application of the main theorem from the previous section along some numerical calculations. All results are obtained for two sets of parameter values —  $\alpha$ ,  $\theta$  and  $\tau$  — of the Burr Type XII distribution. For each set, we consider four probability levels:  $p = 0.005$ ,  $p = 0.05$ ,  $p = 0.995$  and  $p = 0.95$ , representing both low and high quantiles. Additionally, for each set of values, two different  $\varepsilon$  values are considered.

Using Theorem [burr] and the corresponding formula ([burr formula]), for each set of parameter values we calculate the upper bound for the convergence rate of the standard quantile estimator  $\hat{x}_p(n)$  to the true quantile value  $x_p$ , or, equivalently, the lower bound for the appropriate coverage probability. Figure 1–4 summarize these results.

The following two graphs present the coverage probability depending on the sample size, for the Burr Type XII distribution with parameters  $\alpha = 5.1$ ,  $\theta = 590$ ,  $\tau = 4.5$ , focusing on one low quantile ( $x_{0.005}$ ) and one high quantile ( $x_{0.995}$ ), and two fixed values for  $\varepsilon$ .



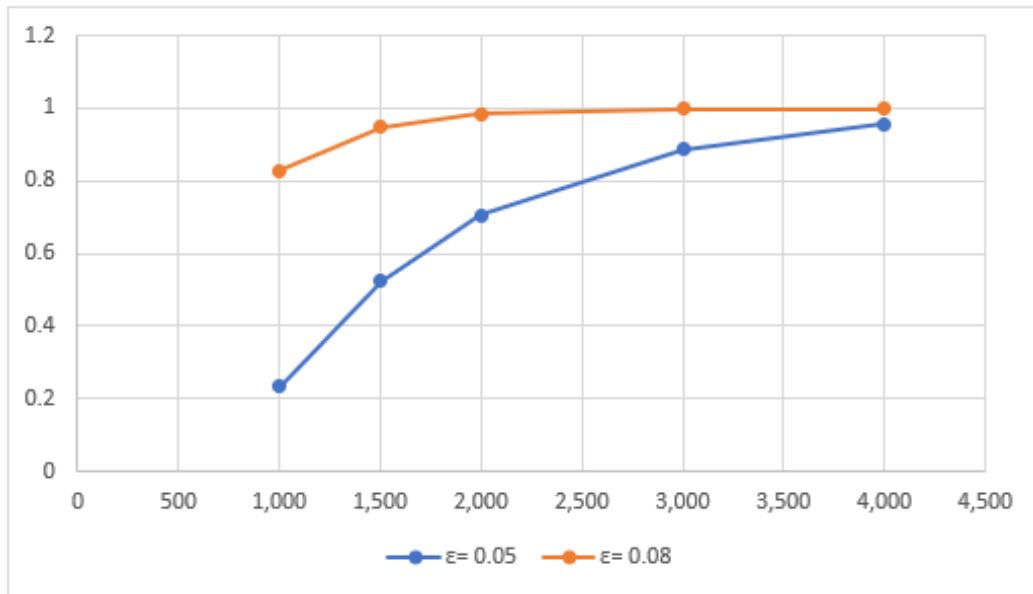
**Figure 1:** Coverage probability vs sample size, for the low quantile with  $p = 0.005$  and  $x_{0.005} = 0.886$



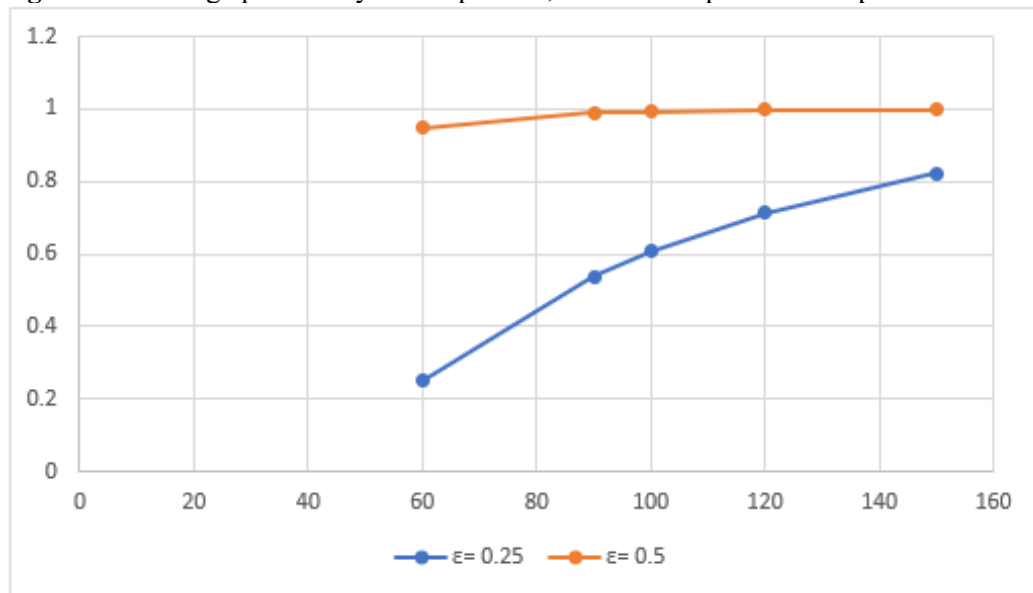
**Figure 2:** Coverage probability vs sample size, for the high quantile with  $p = 0.995$  and  $x_{0.995} = 4.719$

The Figure 1 and 2 show the coverage probability of the standard quantile estimator  $\hat{x}_p(n)$  for the true quantile value  $x_p$ , as a function of sample size and fixed other parameters. It is evident that the coverage probability increases with increasing sample size. For the low quantile  $x_{0.005}$  sample sizes ranging from 2,000 to 15,000 provide appropriate results, while for the high quantile  $x_{0.995}$ , suitable results are obtained with sample sizes between 20 to 110, for considered two cases.

The following two graphs present the coverage probability depending on the sample size, for the Burr Type XII distribution with parameters  $\alpha = 46$ ,  $\theta = 470$ ,  $\tau = 3.2$ , focusing on one low quantile ( $x_{0.05}$ ) and one high quantile ( $x_{0.95}$ ), and two fixed values for  $\epsilon$ .



**Figure 3:** Coverage probability vs sample size, for the low quantile with  $p = 0.05$  and  $x_{0.05} = 0.817$



**Figure 4:** Coverage probability vs sample size, for the high quantile with  $p = 0.95$  and  $x_{0.95} = 2.943$

The Figure 3 and 4 show the coverage probability of the standard quantile estimator  $\hat{x}_p(n)$  for the true quantile value  $x_p$ , as a function of sample size and fixed other parameters. It is evident that the coverage probability increases with increasing sample size. For the low quantile  $x_{0.05}$  sample sizes ranging from 1,000 to 4,000 provide appropriate results, while for the high quantile  $x_{0.95}$ , suitable results are obtained with sample sizes between 60 to 150, for considered two cases.

## 5. CONCLUSION

This study provides both a theoretical framework and empirical evidence for evaluating the coverage probability of the standard quantile estimator under the Burr Type XII distribution, and it is conducted under the assumptions of i.i.d. between observations; however the results may be used in the case of negatively dependent variables as well. By applying large deviations theory, we derived upper bound for the convergence rate, that describe the rate at which the estimator converges to the true quantile. The results confirmed that convergence improves with increasing sample size, in the sense that coverage probability increase with increasing sample size, and is influenced by the shape of the distribution and the quantile level. The Burr Type XII model demonstrated strong performance across parameter configurations that reflect realistic insurance loss scenarios, reinforcing its practical relevance. These results offer valuable insights for improving the reliability of quantile-based risk measures such as Value-at-Risk. Future research may explore similar bounds for alternative estimators or dependence structures, thereby broadening the applicability of this approach in actuarial and financial practice.

## LITERATURE

- Beirlant, J., Matthys, G., & Dierckx, G. (2001). Heavy-tailed distributions and rating. *ASTIN Bulletin: The Journal of the IAA*, 31(1), 37–58. <https://doi.org/10.2143/AST.31.1.993>
- Burr, I. W. (1942). Cumulative frequency functions. *The Annals of mathematical statistics*, 13(2), 215–232. <https://doi.org/doi:10.1214/aoms/1177731607>
- Cummins, J. D., Dionne, G., McDonald, J. B., & Pritchett, B. M. (1990). Applications of the GB2 family of distributions in modeling insurance loss processes. *Insurance: Mathematics and Economics*, 9(4), 257–272. [https://doi.org/10.1016/0167-6687\(90\)90003-V](https://doi.org/10.1016/0167-6687(90)90003-V)
- Jin, X., & Fu, M. C. (2007). A large deviations analysis of quantile estimation with application to value at risk. Robert H. Smith School of Business.
- Miljković, T., Causey, R., & Jovanović, M. (2022). Assessing the performance of confidence intervals for high quantiles of Burr XII and Inverse Burr mixtures. *Communications in Statistics-Simulation and Computation*, 51(8), 4677–4699. <https://doi.org/10.1080/03610918.2020.1747075>
- Miljković, T., & Grun, B. (2016). Modeling loss data using mixtures of distributions. *Insurance: Mathematics and Economics*, 70, 387–396. <https://doi.org/10.1016/j.insmatheco.2016.06.019>
- Stanojević, J., & Jovović, M. (2016). Coverage probability for one quantile estimator and its application on insurance data. *Bulletin of the international mathematical virtual institute*, 6(2016), 127–134.
- Venter, G. (1983). Transformed beta and gamma distributions and aggregate losses. *Proceedings of the Casualty Actuarial Society*, 70(133), 289–308.