The Journal of Economic Inequality (2023) 21:499–509
https://doi.org/10.1007/s10888-022-09556-4

The shortest confidence interval for the ratio of quantiles of the Dagum distribution

Alina Jędrzejczak[1] · Dorota Pekasiewicz[1] · Wojciech Zieliński[2]

Received: 19 June 2019 / Accepted: 31 October 2022 / Published online: 3 December 2022

Abstract

Jędrzejczak et al. (REVSTAT-Statistical Journal **19**(1), 87–97, 2021) constructed a confidence interval for a ratio of quantiles coming from the Dagum distribution, which is frequently applied as a theoretical model in numerous income distribution analyses. The proposed interval is symmetric with respect to the ratio of sample quantiles, which result may be unsatisfactory in many practical applications. The search for a confidence interval with a smaller length resulted in the derivation of the shortest interval with the ends being asymmetric relative to the ratio of sample quantiles. In the paper, the existence of the shortest confidence interval is shown and the method of obtaining such an interval is presented. The results of the calculations show a reduction in the length of the proposed confidence interval by several percent compared to the symmetrical confidence interval.

Keywords Quantile ratio · Ratio of quintiles · The shortest confidence interval · Dagum distribution

1 Introduction

Among the most common indices used to measure income distribution inequality are the ratios of quantiles including the most popular decile ratios and the quintile ratio, also called

✉ Wojciech Zieliński
wojciech_zielinski@sggw.edu.pl

Alina Jędrzejczak
alina.jedrzejczak@uni.lodz.pl

Dorota Pekasiewicz
dorota.pekasiewicz@uni.lodz.pl

1 University of Łódź, Łódź, Poland

2 Department of Econometrics, Warsaw University of Life Sciences, Nowoursynowska 159, PL-02-767 Warsaw, Poland

S80/S20 ratio (Eurostat Regional Yearbook, 2016).[1] It is calculated as the ratio of the fourth quintile of an income distribution to the first one, i.e. income quintile ratio can be defined as

$$r_{0.2,0.8} = \frac{F^{-1}(0.8)}{F^{-1}(0.2)},$$

where F denotes the cumulative distribution function of income.

Natural point estimators of the ratio $r_{0.2,0.8}$ can easily be obtained as ratios of the corresponding sample quintiles. Statistical properties of such estimators are dependent on the form of quantile estimators which have been implemented (see: Jędrzejczak and Pekasiewicz 2018). In income distribution analysis, the interval estimation of different inequality measures is also considered in order to provide information about how close the point estimate is to the true parameter with the margin of error that can be accepted by social policy-makers. In Greselin and Pasquazzi (2009, 2010), parametric Dagum confidence intervals for Gini and new Zenga inequality measures were derived and compared with non-parametric ones. Jędrzejczak et al. (2021) constructed a confidence interval for the ratios of quantiles, assuming the Dagum distribution (1977) as an income distribution model. Also in this paper we confine ourselves to the Dagum distribution as a probabilistic model of size income distribution. This distribution has widely been applied in income inequality analysis in many countries all over the world, due to its economic foundations and high flexibility to fit observed income distributions in total and in various brakdowns (by region, gender, socio-economic group etc.), what was reported in numerous works (Domański and Jędrzejczak 2002; Kleiber and Kotz 2003; Bandourian et al. 2003; Brzeziński 2013). In Domma et al. (2011), the Authors show that the Dagum distribution may be a competitive model for describing data which include censored observations in actuarial sciences and system reliability theory. It is worth noting the three-parameter Dagum distribution preservers the "parsimony" property of being dependent on only a small number of parameters (Dagum 1977).

Comparing to the previous work of the Authors, in this paper we consider a more general set-up, namely a confidence interval for a ratio of α and β quantiles ($0 < \alpha < \beta < 1$) is taken into regard. It comes down to the following formula

$$r_{\alpha,\beta} = \frac{F^{-1}(\beta)}{F^{-1}(\alpha)}.$$

Jędrzejczak et al. (2021) constructed an equitailed confidence interval, i.e confidence interval for which the risks of underestimation and overestimation are the same i.e $P\{r_{\alpha,\beta} < L\} = P\{r_{\alpha,\beta} > U\}$. Here (L, U) denotes the confidence interval for $r_{\alpha,\beta}$ at the confidence level δ:

$$L = \frac{r^*_{\alpha,\beta} z_{\frac{1+\delta}{2}}(a)}{W\left(r^*_{\alpha,\beta} z_{\frac{1+\delta}{2}}(a) \exp\left(z_{\frac{1+\delta}{2}}(a)\right)\right)}, \quad U = \frac{r^*_{\alpha,\beta} z_{\frac{1-\delta}{2}}(a)}{W\left(r^*_{\alpha,\beta} z_{\gamma}(a) \exp\left(z_{\frac{1-\delta}{2}}(a)\right)\right)}$$

(for details see Jędrzejczak et al. 2021). Such a confidence interval will be referred to as the standard one. The current study is dedicated to the problem of construction of the shortest confidence interval.

The second section briefly introduces the Dagum distribution while in the third section the shortest confidence interval is derived. Unfortunately, closed formulae turned out to be

[1] In Eurostat publications the term S20/S80 stands for income quintile share ratio, which is the proportion held by the upper 20% divided by the proportion held by the bottom 20%

not available. Some numerical results are given in the fourth section. In the last section conclusions and final remarks are presented, as well as the suggestions of future research topics

2 The Dagum distribution

The Dagum distribution is often used in the analysis of personal (or household) income and wages, as it is usually well fitted to empirical distributions in different countries. It can also be successfully applied for different subpopulations obtained by means of splitting up the overall sample by socio-economic group, region, gender or family type (Jędrzejczak and Pekasiewicz 2018; Pekasiewicz and Jędrzejczak 2017). The estimates of the Dagum parameters are utilized to assess many important statistical characteristics, including numerous variability, inequality, poverty and wealth measures, as well as concentration curves.

Consider a Dagum distribution with parameters $a > 0$, $v > 0$ and $\lambda > 0$. Its cumulative distribution distribution (CDF) and probability density function (PDF) are as follows

$$F_{a,v,\lambda}(x) = \left(1 + \left(\frac{x}{\lambda}\right)^{-v}\right)^{-a} \quad \text{for } x > 0$$

and

$$f_{a,v,\lambda}(x) = \frac{av}{\lambda}\left(\frac{x}{\lambda}\right)^{av-1}\left(1 + \left(\frac{x}{\lambda}\right)^{v}\right)^{-a-1} \quad \text{for } x > 0.$$

Its quantile function equals

$$Q_{a,v,\lambda}(q) = \lambda\left(q^{-1/a} - 1\right)^{-1/v} \quad \text{for } 0 < q < 1.$$

3 The shortest confidence interval

Among the many characteristics of income inequality, quantile-based measures play an important role. Simple dispersion ratios, defined as the ratios of the income of the richest quantile over that of the poorest quantile, usually utilize deciles and quintiles, but in principle, any quantile of income distribution can be used. A version of the decile dispersion ratio, which has recently become popular, also called the Palma Ratio (Palma 2011), is based on the ratio of the 90th over the 40th percentile (or the richest 10% of the population's income share divided by the poorest 40%'s share) . Another popular inequality measure based on quantiles (deciles) is the coefficient of maximum equalisation, also known as the Schutz index or the Pietra ratio. Furthermore one of the point inequality measures proposed by Zenga (1990) and considered in Kleiber and Kotz (2003), is based on the distribution and income quantiles.

Let $0 < \alpha < \beta < 1$ be given numbers. We are interested in estimation of the ratio of quantiles

$$r_{\alpha,\beta} = \frac{Q_{a,v,\lambda}(\beta)}{Q_{a,v,\lambda}(\alpha)} = \frac{\left(\beta^{-1/a} - 1\right)^{-1/v}}{\left(\alpha^{-1/a} - 1\right)^{-1/v}}.$$

Since we are interested in the estimation of the ratio $r_{\alpha,\beta}$ of quantiles, we reparametrize the considered model. It can be seen that

$$v = \frac{\log\left(\frac{\alpha^{-1/a}-1}{\beta^{-1/a}-1}\right)}{\log r_{\alpha,\beta}}.$$

After appropriate substitution, the CDF of the Dagum distribution may be written in the following form

$$F_{a,r_{\alpha,\beta},\lambda}(x) = \left(1 + \left(\frac{x}{\lambda}\right)^{-\frac{\log\left(\frac{\alpha^{-1/a}-1}{\beta^{-1/a}-1}\right)}{\log r_{\alpha,\beta}}}\right)^{-a}$$

for $x > 0$ and $a > 0$, $r_{\alpha,\beta} > 0$ and $\lambda > 0$.

Let $X_1, \ldots, X_n$ be a simple random sample drawn from a Dagum distribution and let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ be the corresponding order statistics. Let $r^*_{\alpha,\beta} = \frac{X_{(\lfloor \beta n \rfloor+1):n}}{X_{(\lfloor \alpha n \rfloor+1):n}}$ (here $\lfloor x \rfloor$ denotes the greatest integer not greater than x) be an observed quantile ratio. It is assumed that the sample size n is large, i.e. it is assumed that $n \to \infty$.

From David and Nagaraja (2003) and Serfling (1980) it follows that $r^*_{\alpha,\beta}$ is a strongly consistent estimator of $r_{\alpha,\beta}$, for all a, v, λ. Also, it follows (Serfling 1980, th. 2.3.3; David and Nagaraja 2003, th. 10.3 and application of Delta method, see e.g. Greene 2003, p. 913) that for $0 < \alpha < \beta < 1$ the estimator $r^*_{\alpha,\beta}$ is asymptotically normally distributed, i.e.

$$\sqrt{n}\left(r^*_{\alpha,\beta} - r_{\alpha,\beta}\right) \to r_{\alpha,\beta} N\left(0, \left(\log r_{\alpha,\beta}\right)^2 w^2(a)\right),$$

where

$$w^2(a) = \left(\frac{1}{a \log\left(\frac{\alpha^{-1/a}-1}{\beta^{-1/a}-1}\right)}\right)^2 \cdot \left(\frac{1-\beta}{\beta}\frac{1}{(1-\beta^{\frac{1}{a}})^2} + \frac{1-\alpha}{\alpha}\frac{1}{(1-\alpha^{\frac{1}{a}})^2} - 2\frac{1-\beta}{\beta}\frac{1}{(1-\alpha^{\frac{1}{a}})(1-\beta^{\frac{1}{a}})}\right)$$

(for theoretical details see Jędrzejczak et al. 2021). Note that, since $a > 0$ and $0 < \alpha < \beta < 1$ the variance $w^2(a)$ is always finite.

Let δ be a given confidence level. We have (the scale parameter λ is omitted)

$$P_{r,a}\left\{u_{\delta_1-\delta} \leq \sqrt{n}\frac{r^*_{\alpha,\beta} - r_{\alpha,\beta}}{r_{\alpha,\beta}\log r_{\alpha,\beta} w(a)} \leq u_{\delta_1}\right\} = \delta,$$

where $\delta \leq \delta_1 \leq 1$ and u_γ is the γ-quantile of $N(0,1)$ distribution.

Let

$$EoCI(\gamma) = \frac{r^*_{\alpha,\beta} z_\gamma(a)}{W\left(r^*_{\alpha,\beta} z_\gamma(a) \exp\left(z_\gamma(a)\right)\right)},$$

where $z_\gamma(a) = \frac{\sqrt{n}}{u_\gamma w(a)}$ and $W(\cdot)$ is the Lambert W function (see Appendix 2). The confidence interval for $r_{\alpha,\beta}$ at the confidence level δ has the form

$$\left(EoCI\left(\delta_1\right);\ EoCI\left(\delta_1 - \delta\right)\right).$$

The confidence interval with $\delta_1 = (1+\delta)/2$ is the standard (i.e. symmetric) one:

$$\left(EoCI\left(\frac{1+\delta}{2}\right);\ EoCI\left(\frac{1-\delta}{2}\right)\right).$$

The length of the confidence interval is a function of δ_1:

$$L\left(\delta_1\right) = EoCI\left(\delta_1 - \delta\right) - EoCI\left(\delta_1\right).$$

We want to minimize $L\left(\delta_1\right)$ with respect to δ_1.

Table 1 Simulated coverage probabilities based on 10000 repetitions for standard c.i. (confidence level 0.95) from the Dagum distribution

	$r_{0.2,0.8}$				$r_{0.1,0.9}$		
a	2	4	6	a	2	4	6
0.1	0.9488	0.9523	0.9531	0.1	0.9459	0.9468	0.9439
0.5	0.9485	0.9512	0.9492	0.5	0.9526	0.9530	0.9516
1.0	0.9510	0.9547	0.9553	1.0	0.9554	0.9555	0.9551
1.5	0.9482	0.9470	0.9490	1.5	0.9535	0.9533	0.9542
2.0	0.9511	0.9455	0.9483	2.0	0.9564	0.9516	0.9524

Lemma $\frac{W(z)}{W(az)}$ *is decreasing for* $a > 1$; *is constant for* $a = 1$; *is increasing for* $0 < a < 1$.

Proof To obtain the thesis it is enough to observe that function $W(z)$ is increasing for $z > e^{-1}$ and it is convex. Those properties as well as the other interesting properties of the Lambert W function may be found in Corless et al. (1996). □

Theorem *There exists* δ_1 *which minimizes* $L(\delta_1)$.

Proof If $\gamma \in (0.5, 1)$ increases then $z_\gamma \exp(z_\gamma)$ decreases. Since $r^*_{\alpha,\beta} > 1$ and $z = W(ze^z)$ hence $EoCI(\gamma)$ decreases. We have:

if $\delta_1 \searrow \delta$ then $EoCI(\delta_1) \to EoCI(\delta) < \infty$ and $EoCI(\delta_1 - \delta) \nearrow +\infty$;

if $\delta_1 \nearrow 1$ then $EoCI(\delta_1) \searrow -\infty$ and $EoCI(\delta_1 - \delta) \to EoCI(1 - \delta) < \infty$.

Hence

$$\delta_1 \searrow \delta \Rightarrow L(\delta_1) \nearrow +\infty \text{ and } \delta_1 \nearrow 1 \Rightarrow L(\delta_1) \nearrow +\infty.$$

From continuity of $EoCI(\cdot)$ we obtain the thesis. □

Note that for $\delta_1 = \delta$ and $\delta_1 = 1$ we obtain one-sided confidence intervals.

4 Numerical results

The analytical form of δ_1 minimizing the length of the confidence interval for quantile ratios of the Dagum distribution is unavailable but its value can be found numerically. In Appendix 1 there is given a short code in R-project language for finding the minimal-length confidence interval.

Some numerical results are given in tables below for $n = 1000$ and $\delta = 0.95$. Due to the simulation results given in Zieliński et al. (2018) it may be assumed that the sample of size 1000 is large enough to do asymptotics.

Simulated coverage probabilities based on 10000 repetitions are given in Tables 1 (c.i. standard) and 2 (c.i. short). The values of S80/S20 here chosen for the simulations are within the typical range for income distributions, e.g. those observed for EU countries, so they are empirically relevant (https://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do). Having got the values of α and value of $r_{0.2,0.8}$, the value of v can easily be calculated from the quantile function.

Table 2 Simulated coverage probabilities based on 10000 repetitions for short c.i. (confidence level 0.95) from the Dagum distribution

	$r_{0.2,0.8}$				$r_{0.1,0.9}$		
a	2	4	6	a	2	4	6
0.1	0.9486	0.9518	0.9527	0.1	0.9459	0.9471	0.9441
0.5	0.9495	0.9514	0.9486	0.5	0.9525	0.9533	0.9524
1.0	0.9510	0.9552	0.9551	1.0	0.9547	0.9549	0.9547
1.5	0.9482	0.9466	0.9490	1.5	0.9532	0.9526	0.9536
2.0	0.9507	0.9456	0.9483	2.0	0.9557	0.9523	0.9524

For the other values of a, $r_{0.2,0.8}$ and $r_{0.1,0.9}$ results are similar, hence we do not present them here. Our simulations show that for the proposed confidence interval the empirical coverage probability equals the assumed confidence level.

Tables 3 and 4 summarize comparison between standard confidence interval and the shortest one. Table 3 presents the results obtained for $\alpha = 0.2 = 1 - \beta$, while Table 4 presents the results for $\alpha = 0.1 = 1 - \beta$. The respective interval lengths are listed in the "short" and "standard" columns. The last column contains the corresponding length reductions which can be considered the precision gains obtained by means of the proposed estimation method. Note that $1 - \delta_1$ is the risk of underestimation while $\delta_1 - \delta$ is the risk of overestimation (for standard confidence interval both probabilities are equal to $(1 - \delta)/2 = 0.025$).

Note that the results given in Tables 3 and 4 are calculated, not simulated.

5 Example of application

In this section, we would like to demonstrate the possible benefits of using the shortest confidence interval for the ratio of quantiles of the Dagum distribution. On the basis of real world data we constructed confidence intervals for the quintiles and deciles ratios, which may be used to study income distributions inequality. The aim was to show how the shortest c.i. compares to the standard c.i. We conducted the Monte Carlo simulation, regarding the distribution of the estimator of the quantile ratio, for which the confidence interval has been proposed. The basis for the calculations was the sample coming from the Polish Household Budget Survey conducted by the Statistics Poland in 2017.

In the first step, based on the sample of $n = 1519$ households of farmers, the empirical distribution of monthly equivalent income was approximated by means of the Dagum distribution (Fig. 1). The obtained values (ML estimates) of the parameters are: $a = 0.69676$, $v = 2.190266$ and $\lambda = 3.04099$.

The goodness-of-fit of the empirical distribution with the Dagum distribution was evaluated determining the structure similarity coefficient and the Mortara index.

The similarity coefficient is given by the formula (see, e.g., Vielrose 1960):

$$W = \sum_{j=1}^{k} \min(w_j, \hat{w}_j)$$

Table 3 Standard c.i. vs. short c.i.: results calculated (confidence level 0.95) for the Dagum distribution for $\alpha = 0.2 = 1 - \beta$

a	$r_{0.2,0.8}$	$\delta_1 - \delta$	$1 - \delta_1$	Short	Standard	Reduction
0.1	2	0.03339	0.01661	0.241322	0.244047	1.117%
0.1	4	0.03646	0.01354	0.978297	1.000800	2.248%
0.1	6	0.03815	0.01185	1.915430	1.976760	3.103%
0.5	2	0.03217	0.01783	0.202939	0.204560	0.792%
0.5	4	0.03486	0.01514	0.819397	0.832625	1.589%
0.5	6	0.03633	0.01367	1.599460	1.635210	2.186%
1.0	2	0.03176	0.01824	0.191330	0.192689	0.705%
1.0	4	0.03436	0.01564	0.771701	0.782753	1.412%
1.0	6	0.03577	0.01423	1.505170	1.534970	1.942%
1.5	2	0.03176	0.01824	0.190684	0.192029	0.700%
1.5	4	0.03430	0.01570	0.769049	0.779988	1.402%
1.5	6	0.03577	0.01423	1.499930	1.529430	1.929%
2.0	2	0.03176	0.01824	0.191483	0.192845	0.706%
2.0	4	0.03436	0.01564	0.772328	0.783407	1.414%
2.0	6	0.03581	0.01420	1.506410	1.536280	1.945%

where k is the number of intervals in which the individual values are grouped, w_j are the observed relative frequencies and $\hat{w}_j$ are the estimated relative frequencies.

Table 4 Standard c.i. vs. short c.i.: results calculated (confidence level 0.95) for the Dagum distribution for $\alpha = 0.1 = 1 - \beta$

a	$r_{0.1,0.9}$	$\delta_1 - \delta$	$1 - \delta_1$	Short	Standard	Reduction
0.1	2	0.03308	0.01692	0.231886	0.234303	1.032%
0.1	4	0.03608	0.01392	0.939046	0.958947	2.075%
0.1	6	0.03768	0.01232	1.837100	1.891220	2.862%
0.5	2	0.03176	0.01824	0.190784	0.192131	0.701%
0.5	4	0.03436	0.01564	0.769460	0.780416	1.404%
0.5	6	0.03577	0.01423	1.500740	1.530209	1.931%
1.0	2	0.03120	0.01880	0.174992	0.176031	0.591%
1.0	4	0.03364	0.01636	0.704826	0.713249	1.181%
1.0	6	0.03496	0.01504	1.373320	1.395970	1.623%
1.5	2	0.03116	0.01884	0.173367	0.174378	0.580%
1.5	4	0.03354	0.01646	0.698190	0.706377	1.159%
1.5	6	0.03486	0.01514	1.360260	1.382270	1.592%
2.0	2	0.03120	0.01880	0.174117	0.175141	0.585%
2.0	4	0.03358	0.01642	0.701252	0.709548	1.169%
2.0	6	0.03490	0.01510	1.366290	1.388600	1.606%

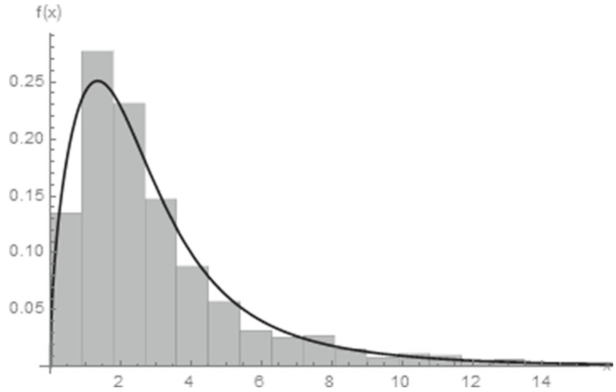

Fig. 1 Empirical income distribution of the Polish farmers and its approximation based on the Dagum model

The Mortara index has the following form (see, e.g., Zenga et al. 2012):

$$IM = \frac{1}{n}\sum_{j=1}^{k}|n_j - \hat{n}_j|$$

where n_j are the observed frequencies of the j-th interval and $\hat{n}_j$ are the estimated frequencies of the j-th interval.

As a result of the calculations, the following values were obtained: $W = 0.9707$ and $IM = 0.0580$ (the observed range was devided into $k = 30$ intervals of the same width.)

The values of consistency measures prove that the Dagum distribution is well fitted to the empirical distribution of farmers' households income for the year 2017. This is also confirmed by the relative difference between the sample mean and the expected value of the Dagum distribution which is as small as 1.5959%.

In the second step, the parameters obtained at the first step were utilized to generate $n = 1000$ element random samples from the Dagum model. In the next step, for each random sample, we evaluated the quantile ratios, namely the quintile and decile ratios.

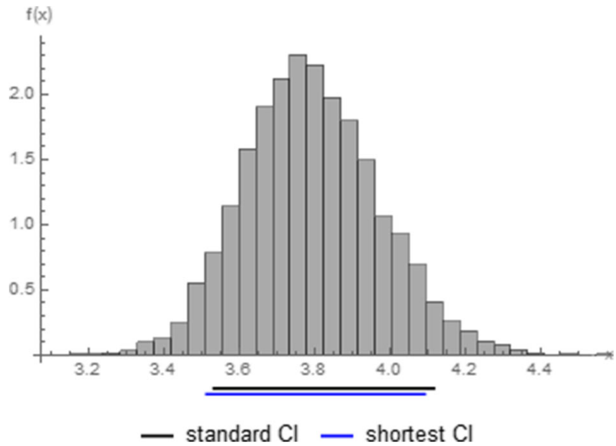

Fig. 2 Empirical (simulated) distribution of the quintile ratio of the Dagum model for $n = 1000$, $N = 10000$

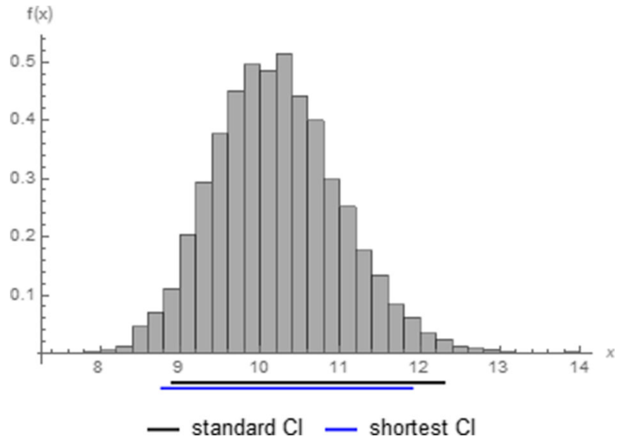

Fig. 3 Empirical (simulated) distribution of the decile ratio of the Dagum model for $n = 1000$, $N = 10000$

Repeating the random sampling and calculating the empirical quantile ratios $N = 10000$ times, a simulated distribution for the quantile ratio has been obtained. Finally, drawing the histogram of each simulated quantile ratio, the Lower and Upper ends of the confidence intervals have been highlighted in the plot, comparing the standard and the shortest c.i. (Figs. 2 and 3).

In Fig. 2 we present the simulated distribution of the ratio of 0.8 and 0.2 quintiles, based on $N = 10000$ repetitions, along with the standard and shortest confidence intervals, estimated at the confidence level of 0.95. The standard c.i. is (3.53323, 4.11604) and the shortest one is (3.51405, 4.09156).

Figure 3 shows the simulated distribution of the ratio of 0.9 and 0.1 deciles, based on $N = 10000$ repetitions, along with the standard and shortest confidence intervals, estimated at the confidence level of 0.95. The standard c.i. is (8.90589, 12.1299) and the shortest one is (8.76803, 11.9124).

It is worth noting that the lower and upper limits of the shortest confidence intervals are in both cases shifted to the left from the limits of the respective standard intervals, which are forced to the right tail of the distribution by extremely high incomes. The efficiency gains offered by the proposed method may be important for social decision-makers who would like to assess the impact of policies, for example programs aimed at combating inequality, in a relatively short time perspective when differences in observed inequality indicators may be still small.

6 Conclusions

One of the crucial problems in socio-economic research is estimation of income inequality which can be evaluated, among others, by the ratio of appropriate quantiles of an income distribution. Such an approach is very convenient for practitioners, as the inequality measures based on quantiles are easy to obtain and have straightforward economic interpretation. Moreover, these measures enable the assessment of inequalities in the extreme parts of the distribution, which is a perfect complement to popular synthetic inequality indices and meets the current challenges of increasing polarization. Nonetheless, reliable economic policy decisions can only be made with the margin of error which should be as small as possible. This requirement meets the proposed confidence interval.

In this paper we proposed the shortest confidence interval. We confined ourselves to the Dagum distribution which was assumed as an underlying income distribution model throughout the paper. It was just because this distribution presents statistical properties required for a good income distribution model and is widely applied in numerous empirical analyses. The confidence interval we constructed is asymptotic, however note that in the real-world experiments on income and wage distributions thousands of data are available. Numerous simulation studies performed in Zieliński et al. (2018) revealed that under the Dagum model the sample size $n = 1000$ is large enough to apply Central Limit Theorem.

The empirical analysis of the lengths of c.i. for quintile and decile ratios confirmed a reduction in the length of the proposed confidence interval by several percent with respect to the symmetric one. It is worth noting that the observed length reduction has strictly been related to the statistical characteristics of the Dagum distribution, namely its dispersion and inequality. The greater income inequality is observed the smaller the precision of interval estimation can be expected and the more reduction you can get due to the new approach. Therefore, the proposed shortest confidence interval can be applied in various income, wage and expenditure analysis, wherever we can successfully utilize the Dagum distribution. Because nowadays it is easy to calculate the shortest confidence intervals hence these intervals can be recommended for practical use. In the future, further investigations on confidence intervals for income inequality and poverty measures, involving different probability distributions, seem useful.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10888-022-09556-4.

Data Availability The data that support the findings of this study are available from the corresponding author upon request.

Declarations

Conflict of Interests All authors declare no conflict of interest.

References

Bandourian, R., McDonald, J.B., Turley, R.S.: A comparison of parametric models of income distribution across countries and over time. Estadistica **55**, 135–152 (2003)

Brzeziński, M.: Parametric modelling of income distribution in Central and Eastern Europe. Cent. Eur. J. Econ. Model. Econ. (CEJEME) **5**(3), 207–230 (2013)

Corless, R.M., Gonnet, G.H., Hare, D.E.G., Jeffrey, D.J., Knuth, D.E.: On the Lambert *W* function. Adv. Comput. Math. **5**(1), 329–359 (1996)

Dagum, C.: A new model of personal income distribution: specification and estimation. Econ. Appl. **30**, 413–437 (1977)

David, H.A., Nagaraja, H.N.: Order Statistics, 3rd edn. Wiley, New York (2003)

Domański, C., Jędrzejczak, A.: Income inequality analysis in the period of economic transformation in Poland. Int. Adv. Econ. Res. **8**(3), 215–222 (2002)
Domma, F., Giordano, S., Zenga, M.M.: Maximum likelihood estimation in Dagum distribution with censored sample. J. Appl. Stat. **38**, 2971–2985 (2011)
Eurostat: The Eurostat regional yearbook ISBN: 978-92-79-60090-6, ISSN: 2363$\bar{1}$716. https://doi.org/10.2785/29084. cat. number: KS-HA-16-001-EN-N. (http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Income_quintile_share_ratio) (2016)
Greene, W.H.: Econometric Analysis, 5th edn. Prentice Hall, Upper Saddle River (2003)
Greselin, F., Pasquazzi, L.: Asymptotic confidence intervals for a new inequality measure. Commun. Stat. Simul. Comput. **38**(8), 1742–1756 (2009)
Greselin, F., Pasquazzi, L.: Dagum confidence intervals for inequality measures. In: Proceedings of the 45th Scientific Meeting of the Italian Statistical Society. Cleup Eds., Padua (2010). https://www.academia.edu/17707359/Dagumconfidenceintervalsforinequalitymeasures
Jędrzejczak, A., Pekasiewicz, D.: Properties of selected inequality measures based on Quantiles and their application to the analysis of income distribution in Poland by macroregion. Argumenta Oeconomica Cracoviensia **18**, 51–67 (2018)
Jędrzejczak, A., Pekasiewicz, D., Zieliński, W.: Confidence interval for Quantile ratio of the Dagum distribution. REVSTAT-Stat. J. **19**(1), 87–97 (2021)
Kleiber, C., Kotz, S.: Statistical Size Distributions in Economics and Actuarial Sciences. Wiley, Hoboken (2003)
Palma, J.G.: Homogeneous middles vs heterogeneous tails, and the end of the 'Inverted-U': the share of the rich is what it's all about. Cambridge Working Papers in Economics (CWPE):1111 (2011)
Pekasiewicz, D., Jędrzejczak, A.: Application of the measures based on quantiles to the analysis of income inequality and poverty in Poland by socio-economic group. In: 35th International Conference Mathematical Methods in Economics, pp. 532–537 (2017)
Serfling, R.J.: Approximation Theorems of Mathematical Statistics. Wiley, New York (1980)
Vielrose, E.: Income distribution by size. Polskie Wydawnictwo Gospodarcze, Warszawa (1960). (in Polish)
Zenga, M.: Concentration curves and concentration indexes derived from them. In: Dagum, C., Zenga, M. (eds.) Income and Wealth Distribution, Inequality and Poverty, pp. 94–110. Springer, Berlin (1990)
Zenga, M.M., Pasquazzi, L., Zenga, M.: First applications of a new three-parameter distribution for non-negative variables. Statistica & Applicazioni **IX**(1), 5–33 (2012)
Zieliński, W., Jędrzejczak, A., Pekasiewicz, D.: Estimation of quantile ratios of the Dagum distribution. In: The 12th Professor Aleksander Zeliaś International Conference on Modelling and Forecasting of Socio-economic Phenomena, vol. 12, pp. 603–611 (2018). https://doi.org/10.14659/SEMF.2018.01.61