LQ-Moments: An Alternative to Standard L- and TL-Moments

DIANA BÍLKOVÁ
Department of Statistics and Probability
University of Economics, Prague
Faculty of Informatics and Statistics
Sq, W. Churchill 1938/4, 130 67 Peague 3
CZECH REPUBLIC
bilkova@vse.cz https://kstp.vse.cz/o-katedre/clenove-katedry/diana-bilkova/

Abstract: The method of LQ-moments represents one of many alternatives to established population parameter estimation techniques. It is used in applied research in such fields as construction, meteorology or hydrology. The present paper focuses on the use of LQ-moments in economics, specifically in wage distribution modelling. The aim of the study is to highlight the advantages of this approach over other methods of estimating the parameters of continuous probability distributions (i.e. those of L- and TL-moments), the theoretical probability distribution being represented by three-parameter lognormal curves. The calculation of sample LQ-moments and identification of the statistical characteristics (those of level, variability, skewness and kurtosis) of a continuous probability distribution are also an integral part of the paper.

Key-Words: LQ-moments, probability distribution, normal distribution, lognormal distribution, large sample theory, comparison with L-moments and TL-moments

1 Introduction

The point estimation of parameters remains a widely discussed issue in the statistical literature, linear quantile (LQ) moments representing a more robust alternative to well-established methods of linear (L) and trimmed linear (TL) moments. Mudholkar & Hutson (1998), for example, introduce LQ-moments as analogues of L-moments acquired by replacing the expectations by functionals inducing the median, Gastwirth estimator and trimean. The same estimators are dealt with by Shabri & Jemain (2006a, 2006b) who develop extended and improved class of LQ-moments that do not impose restrictions on the values of p and α, their combinations lying within the range 0–0.5. Respectively, they design a weighted kernel estimator for quantile function estimation and conduct Monte Carlo simulations to check the performance of the proposed estimators of the three-parameter lognormal distribution. Shabri & Jemain (2010) also adapt the method of LQ-moments for a four-parameter kappa distribution considered as a combination of generalized distributions. Šimková & Picek (2017) derive L-, LQ- and TL-moments of generalized Pareto and extreme-value distributions up to the fourth order, using the first three moments to obtain estimators of their parameters. Performing a simulation study, they compare high-quantile estimates based on L-, LQ-, and TL-moments with the maximum likelihood estimate in terms of their respective sample mean squared errors. Ashour, El-Sheik & Abu El-Magd (2015) derive both TL-and LQ-moments of the exponentiated Pareto distribution, applying them to estimate the unknown parameters. In addition to distribution classification and model selection criteria, Mudholkar & Natarajan (2002) deal with L and LQ measures of skewness and kurtosis, noting that the former measures occur only in the case of finite expectation distributions. David & Nagaraja (2003) also consider measurements of probability distributions and some quick parameter estimators. Abu El-Magd (2010) obtains TL- and LQ-moments of the exponentiated generalized extreme value distribution and utilizes them to estimate the unknown parameters, dealing with some specific cases such as L-, LH- and LL-moments. Deka, Borah & Kakaty, (2009) determine the best-fitting distribution to describe annual time series of maximum daily rainfall data from nine measuring stations in North-East India for the period 1966–2007. GEV, generalized logistic, generalized Pareto, lognormal and Pearson distributions are fitted for this purpose employing L- and LQ-moments. Zin, Jemain & Ibrahim (2009) and Bhuyan & Borah (2011) use the same methodology and the five above-mentioned probability distributions, the former researchers finding the best fitting distribution to analyse annual series of rainfall measured on the Malaysian peninsula between 1975 and 2004. Zaher, El-Sheik & Abu El-Magd (2014) obtain TL-, L- and LQ-moments formulas for the Pareto distribution, comparing the fuzzy least-squares estimator for the two-parameter distribution with other types of estimators. Parameter estimation using LQ-moments is the basis of research in quantile models for Zin & Jemain (2008), who apply thirteen methods of non-parametric quantile estimation to six types of extreme distributions, assessing their effectiveness. A general overview of recent research on L-, Tl- and LQ-moments is provided in Kandeel (2015).

2 LQ-Moments

The method of L-moments – linear functions of the expected values in order statistics – has been widely used in different areas of applied research such as construction, hydrology and meteorology. Its LQ variant is addressed in the present study. LQ-moments are obtained by substituting the expected values with functionals inducing quick estimators such as the median, trimean and Gastwirth mean. They are easy to estimate and evaluate. Skewness and kurtosis measurements based on LQ-moments represent more appropriate and efficient alternatives to standard beta coefficients, the asymptotic distribution of LQ estimators proving their effective simplicity. Their application, particularly in hydrological analysis of extreme flood data values, is discussed in the statistical literature. Other potential uses are outlined in this paper.

2.1 LQ-Moments of Probability Distribution

Let X_1, X_2, …, X_n be a random sample from a continuous distribution with distribution and quantile functions $F_X(\cdot)$ and $Q_X(u) = F_X^{-1}(u)$, respectively, $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$ representing order statistics. Then the r-th L-moment λ_r is given as

$$\lambda_r = r^{-1} \cdot \sum_{k=0}^{r-1} (-1)^k \cdot \binom{r-1}{k} \cdot E(X_{r-k:r}), \quad r = 1, 2, \ldots . \tag{1}$$

Analogously, we define the r-th LQ-moment ξ_r as

$$\xi_r = r^{-1} \cdot \sum_{k=0}^{r-1} (-1)^k \cdot \binom{r-1}{k} \cdot \tau_{p,\alpha}(X_{r-k:r}), \quad r = 1, 2, \ldots , \tag{2}$$

where $0 \leq \alpha \leq ½$, $0 \leq p \leq ½$ and

$$\tau_{p,\alpha}(X_{r-k:r}) = pQ_{X_{r-k:r}}(\alpha) + (1-2p)Q_{X_{r-k:r}}(1/2) + pQ_{X_{r-k:r}}(1-\alpha) . \tag{3}$$

It is evident from equations (1) and (2) that the expected value $E(\cdot)$ at point $\tau_{p,\alpha}(\cdot)$ in the latter equation defines L-moments. Another generalization of L-moments, including the replacement of the expected value in equation (1), is possible by TL-moments.

The linear combination $\tau_{p,\alpha}$, defined by equation (3) is a quick measure of the level of the random distribution of the order statistics $X_{r-k:r}$. Candidates for $\tau_{p,\alpha}$ include functionals generating common quick estimators –

median $$Q_{X_{r-k:r}}\left(\frac{1}{2}\right), \tag{4}$$

trimean $$Q_{X_{r-k:r}}\left(\frac{1}{4}\right)\Big/4 + Q_{X_{r-k:r}}\left(\frac{1}{2}\right)\Big/2 + Q_{X_{r-k:r}}\left(\frac{3}{4}\right)\Big/4, \tag{5}$$

Gastwirth $$0{,}3Q_{X_{r-k:r}}\left(\frac{1}{3}\right) + 0{,}4Q_{X_{r-k:r}}\left(\frac{1}{2}\right) + 0{,}3Q_{X_{r-k:r}}\left(\frac{2}{3}\right). \tag{6}$$

When sampling from a normal distribution, the LQ-Gastwirth estimator is the most efficient considering the possibilities given by equations (4)−(6). The following four LQ-moments of the random variable X are commonly used in practical applications such as probability density classification and parameter estimation

$$\xi_1 = \tau_{p,\alpha}(X), \tag{7}$$

$$\xi_2 = \frac{1}{2}\left[\tau_{p,\alpha}(X_{2:2}) - \tau_{p,\alpha}(X_{1:2})\right], \tag{8}$$

$$\xi_3 = \frac{1}{3}\left[\tau_{p,\alpha}(X_{3:3}) - 2\,\tau_{p,\alpha}(X_{2:3}) + \tau_{p,\alpha}(X_{1:3})\right], \tag{9}$$

$$\xi_4 = \frac{1}{4}\left[\tau_{p,\alpha}(X_{4:4}) - 3\,\tau_{p,\alpha}(X_{3:4}) + 3\,\tau_{p,\alpha}(X_{2:4}) - \tau_{p,\alpha}(X_{1:4})\right]. \tag{10}$$

It is obvious that location measures $\tau_{p,\alpha}(\cdot)$ exist for any random variable X. Therefore, the r-th LQ-moment always exists, and it is unique if the distribution function $F_X(\cdot)$ is continuous. Moreover, the evaluation of LQ-moments of any continuous distribution can be simplified if the following applies: $Q_X(\cdot) = F_X^{-1}(\cdot)$ being the quantile function of a random variable X, a quick measure of the level defined by equation (3) is equivalent to the equation

$$\tau_{p,\alpha}(X_{r-k:r}) = p\,Q_X[B_{r-k:r}^{-1}(\alpha)] + (1-2p)Q_X[B_{r-k:r}^{-1}(1/2)] + p\,Q_X[B_{r-k:r}^{-1}(1-\alpha)], \tag{11}$$

where $B_{r-k:r}^{-1}(\alpha)$ denotes the corresponding α-th quantile of a beta distributed random variable with parameters $r-k$ and $k+1$.

When constructing appropriate distribution models and estimating parameters, the coefficients of skewness and kurtosis $\sqrt{\beta_1}$ and β_2, respectively, play an important role in terms of the classification of statistical distributions. Due to their drawbacks, however, alternative measures of skewness and kurtosis are used, including relatively recent ones such as

$$\gamma_U(F) = [F^{-1}(1-u) + F^{-1}(u) - 2m_F]/[F^{-1}(1-u) + F^{-1}(u)].$$

Quantile-based measures of kurtosis for symmetric distributions include

$$[Q(0{,}75+u) + Q(0{,}75-u) - 2Q(0{,}75)]/$$

$$/[Q(0{,}75+u) - Q(0{,}75-u)],\ 0 \le u < 1/4$$

and

$$[Q(0{,}5+u) - Q(0{,}5-u)]/[Q(0{,}75) - Q(0{,}25)],\ 0 \le u < 1/2.$$

L-moment-based ratios τ_3 and τ_4 – L-skewness and L-kurtosis, respectively – are defined as

$$\tau_3 = \frac{\lambda_3}{\lambda_2} \quad \text{and} \quad \tau_4 = \frac{\lambda_4}{\lambda_2}, \tag{12}$$

offering an alternative to $\sqrt{\beta_1}$ and β_2. It is proved that τ_3 meets the convex arrangement, τ_4 maintaining van Zwet's symmetric ordering.

The skewness and kurtosis measures η_3 and η_4 based on LQ-moments – LQ-skewness and LQ-kurtosis – are defined as

$$\eta_3 = \frac{\xi_3}{\xi_2} \quad \text{and} \quad \eta_4 = \frac{\xi_4}{\xi_2}. \tag{13}$$

It is necessary to note that both LQ-skewness and LQ-kurtosis exist for all distributions and are invariant in terms of location and scale. However, other analogous properties of τ_3 and τ_4 mentioned above remain unexplored for LQ-skewness η_3 and LQ-kurtosis η_4, their behaviour being now more thoroughly analysed.

Another ratio measurement useful for comparing distributions with the usual origin and scale is an analogy of the coefficient of variation

$$\eta_2 = \frac{\xi_2}{\xi_1}, \tag{14}$$

where ξ_1 and ξ_2 are represented by equations (7) and (8). When modelling survival data, it is common practice to plot $\sqrt{b_1}$ against the sample coefficient of variation

$$\hat{\gamma} = \frac{s}{\bar{x}} \tag{15}$$

in the $(\sqrt{\beta_1}, \gamma)$ plane to verify the model selection.

2.2 Sample LQ-Moments

LQ-moments can be estimated directly by estimating the quantiles of order statistics in combination with equation (11). The simplest quantile estimator suitable for this purpose is the one based on linear interpolation, available in standard statistical software packages. However, alternative estimators of quantiles can be used as well.

Let $X_{1:n} \le X_{2:n} \le \ldots \le X_{n:n}$ be the sample order statistics. The quantile estimator $Q(u)$ is then given by

$$\hat{Q}_X(u) = (1-\varepsilon)\, X_{[n'u]:n} + \varepsilon\, X_{[n'u]+1:n}, \tag{16}$$

where $\varepsilon = n'u - [n'u]$ and $n' = n + 1$.

For random samples of sample size n, the r-th sample LQ-moment is expressed by the relationship

$$\hat{\xi}_r = r^{-1} \cdot \sum_{k=0}^{r-1} (-1)^k \cdot \binom{r-1}{k} \cdot \hat{\tau}_{p,\alpha}(X_{r-k:r}), \quad r = 1, 2, \dots, \tag{17}$$

where $\hat{\tau}_{p,\alpha}(X_{r-k:r})$ is a quick estimator of the distribution of the order statistic $X_{r-k:r}$ in a random sample of size r.

Specifically, the first four sample LQ-moments from equation (17) are given as

$$\hat{\xi}_1 = \hat{\tau}_{p,\alpha}(X), \tag{18}$$

$$\hat{\xi}_2 = \frac{1}{2}\left[\hat{\tau}_{p,\alpha}(X_{2:2}) - \hat{\tau}_{p,\alpha}(X_{1:2})\right], \tag{19}$$

$$\hat{\xi}_3 = \frac{1}{3}\left[\hat{\tau}_{p,\alpha}(X_{3:3}) - 2\,\hat{\tau}_{p,\alpha}(X_{2:3}) + \hat{\tau}_{p,\alpha}(X_{1:3})\right], \tag{20}$$

$$\hat{\xi}_4 = \frac{1}{4}\left[\hat{\tau}_{p,\alpha}(X_{4:4}) - 3\,\hat{\tau}_{p,\alpha}(X_{3:4}) + 3\,\hat{\tau}_{p,\alpha}(X_{2:4}) - \hat{\tau}_{p,\alpha}(X_{1:4})\right], \tag{21}$$

where the quick estimator $\hat{\tau}_{p,\alpha}(X_{r-k:r})$ of the level of order statistics $X_{r-k:r}$ is described by relationship

$$\begin{aligned} \hat{\tau}_{p,\alpha}(X_{r-k:r}) &= p\,\hat{Q}_{X_{r-k:r}}(\alpha) + (1-2p)\,\hat{Q}_{X_{r-k:r}}(1/2) + \\ &\quad + p\,\hat{Q}_{X_{r-k:r}}(1-\alpha), \\ &= p\,\hat{Q}_X[B^{-1}_{r-k:r}(\alpha)] + (1-2p)\,\hat{Q}_X[B^{-1}_{r-k:r}(1/2)] + \\ &\quad + p\,\hat{Q}_X[B^{-1}_{r-k:r}(1-\alpha)], \end{aligned} \tag{22}$$

where $0 \le \alpha \le ½$, $0 \le p \le ½$, $B^{-1}_{r-k:r}(\alpha)$ is the α-th quantile of the random variable with a beta distribution with parameters $r - k$ a $k + 1$, and $\hat{Q}_X(\cdot)$ denotes an estimator using linear interpolation given by equation (16). The calculation of sample LQ-moment $\hat{\xi}_r$ is as follows

$$\hat{\xi}_4 = \frac{1}{4}\left[\hat{\tau}_{p,\alpha}(X_{4:4}) - 3\,\hat{\tau}_{p,\alpha}(X_{3:4}) + 3\,\hat{\tau}_{p,\alpha}(X_{2:4}) - \hat{\tau}_{p,\alpha}(X_{1:4})\right],$$

being simplified using quantile $B^{-1}_{r-k:r}(\alpha)$ that can be easily obtained from statistical spreadsheets.

Explicit schemes for the calculation of LQ-moments are presented, the three quick estimators – median ($p = 0$, $\alpha = \cdot$), trimean ($p = 1/4$, $\alpha = 1/4$) and Gastwirth ($p = 0{,}3$, $\alpha = 1/3$) – are used for $\hat{\tau}_{p,\alpha}(X_{r-k:r})$ given by equation (22). The calculation of the first four sample LQ-moments from equation (17) is simplified using pyramid schemes.

Sample LQ-skewness and LQ-kurtosis

$$\hat{\eta}_3 = \frac{\hat{\xi}_3}{\hat{\xi}_2} \quad \text{a} \quad \hat{\eta}_4 = \frac{\hat{\xi}_4}{\hat{\xi}_2}, \tag{23}$$

can be used to identify η_3 and η_4 and to estimate parameters.

2.3 Large Sample Theory

Sample LQ-moments depend on the choice of quick and quantile estimators, their asymptotic normality, however, being consistent with the theory of linear order statistics of large samples. In order to develop expressions for the large sample mean and variance of sample LQ-moments, we shall limit ourselves to the Q class of quantile functions Q meeting the following conditions:

- the inverse function $Q_X(u) = F_X^{-1}(u)$ is defined exclusively for $0 < u < 1$;
- $Q(\cdot)$ is twice differentiable on the interval (0, 1) with a continuous second derivative $Q''(\cdot)$ on the same interval;
- $Q'(\cdot) > 1$ for $0 < u < 1$.

Let us consider $0 < u_1 < u_2 < \dots < u_k < 1$, assuming the above conditions (1)–(3) are fulfilled. Then

$$[\hat{Q}(u_1), \hat{Q}(u_2), \dots, \hat{Q}(u_k)]$$

is asymptotically normal with a vector of expected values $[Q(u_1), Q(u_2), \dots, Q(u_k)]$ as well as with covariances

$$\begin{aligned} \sigma_{ij} &= Cov[\hat{Q}(U_i), \hat{Q}(U_j)] = \\ &= u_i\,(1-u_j)\,Q'(u_i)\,Q'(u_j)/n, \quad i \le j, \sigma_{ij} = \sigma_{ji}. \end{aligned} \tag{24}$$

To create asymptotic expressions for covariances of LQ-moments, we will first obtain

$$Cov[\hat{\tau}_{p,\alpha}(X_{r-k:r}), \hat{\tau}_{p,\alpha}(X_{s-l:s})],$$

which is a function dependent on six specific percentiles $u_1, u_2, \ldots, u_6$ used to derive equation (24) rewritten into a set of six percentiles

$$B^{-1}_{r-k:r}(\alpha),\ B^{-1}_{s-l:s}(\alpha),\ B^{-1}_{r-k:r}(1/2),\ B^{-1}_{s-l:s}(1/2),$$

$$B^{-1}_{r-k:r}(1-\alpha),\ B^{-1}_{s-l:s}(1-\alpha),$$

so that $0 < u_1 < u_2 < \ldots < u_6 < 1$, where $B^{-1}_{r-k:r}(\alpha)$ represents the α-th quantile of a random variable with a beta distribution with parameters $r-k$ and $k+1$. Then, we can get $Cov[\hat{Q}(U_i), \hat{Q}(U_j)]$.

The covariance between the estimated quick estimators of order statistics is defined as

$$\begin{aligned} &Cov[\hat{\tau}_{p,\alpha}(X_{r-k:r}), \hat{\tau}_{p,\alpha}(X_{s-l:s})] = \\ &= p\,\{p\,Cov[\hat{Q}(u_1), \hat{Q}(u_2)] + (1-2p)\,Cov[\hat{Q}(u_2), \hat{Q}(u_3)] + \\ &+ p\,Cov[\hat{Q}(u_2), \hat{Q}(u_5)] + p\,Cov[\hat{Q}(u_1), \hat{Q}(u_6)] + \\ &+ (1-2p)\,Cov[\hat{Q}(u_3), \hat{Q}(u_6)] + p\,Cov[\hat{Q}(u_5), \hat{Q}(u_6)]\} + \\ &+ (1-2p)\,\{p\,Cov[\hat{Q}(u_1), \hat{Q}(u_4)] + (1-2p)\,Cov[\hat{Q}(u_3), \hat{Q}(u_4)] + \\ &+ p\,Cov[\hat{Q}(u_4), \hat{Q}(u_5)]\}\,. \end{aligned} \quad (25)$$

The r-th sample LQ-moment

$$\hat{\xi}_r,\ r = 1, 2, \ldots,$$

has an asymptotically normal distribution with expected value ξ_r. For $r \le s$, covariances of LQ-moments are given by equation

$$\begin{aligned} Cov(\hat{\xi}_r, \hat{\xi}_s) &= \frac{1}{r\,s}\sum_{k=0}^{r-1}\sum_{l=0}^{s-1}(-1)^{k+l}\binom{r-1}{k}\binom{s-1}{l}\cdot \\ &\cdot Cov[\hat{\tau}_{p,\alpha}(X_{r-k:r}), \hat{\tau}_{p,\alpha}(X_{s-l:s})], \end{aligned} \quad (26)$$

where $Cov[\hat{\tau}_{p,\alpha}(X_{r-k:r}), \hat{\tau}_{p,\alpha}(X_{s-l:s})]$ is described by equation (25) and $u_1, u_2, \ldots, u_6$ are specified above. For $r = s$, we obtain the variance of the r-th sample LQ-moment

$$\hat{\xi}_r.$$

As $n \to \infty$, sample measures of LQ-skewness

$$\hat{\eta}_3$$

and LQ-kurtosis

$$\hat{\eta}_4$$

have a two-dimensional normal distribution with the vector of expected values (η_3, η_4) and

$$Var(\hat{\eta}_3) = Var(\hat{\xi}_3)/\hat{\xi}_2^2, \quad (27)$$

$$\begin{aligned} Cov(\hat{\eta}_3, \hat{\eta}_4) &= [Cov(\hat{\xi}_3, \hat{\xi}_4) - \hat{\xi}_3\,Cov(\hat{\xi}_2\ \hat{\xi}_4) - \\ &- \hat{\xi}_4\,Cov(\hat{\xi}_2, \hat{\xi}_3) + \hat{\xi}_3\hat{\xi}_4\,Var(\hat{\xi}_2)]/\hat{\xi}_2^2, \end{aligned} \quad (28)$$

$$Var(\hat{\eta}_4) = Var(\hat{\xi}_4)/\hat{\xi}_2^2, \quad (29)$$

where

$$Var(\hat{\xi}_r) = Cov(\hat{\xi}_r, \hat{\xi}_r)$$

and variances and covariances indicate the right side of the equation (26).

2.4 Application to Normal Distribution

We consider a random sample from a normal distribution and compare the use of the median, trimean and Gastwirth estimators when estimating LQ-skewness and LQ-kurtosis. Then the estimators

$$\hat{\eta}_3$$

and

$$\hat{\eta}_4$$

given by equation (23) have a common normal distribution with the corresponding expected value vectors

$$(0;\ 0{,}116),\ (0;\ 0{,}118) \text{ and } (0;\ 0{,}117) \quad (30)$$

and covariance matrices

$$\mathbf{\Sigma}_{\text{MED}} = \frac{1}{n}\begin{pmatrix} 1{,}535 & 0 \\ 0 & 2{,}070 \end{pmatrix}, \mathbf{\Sigma}_{\text{TRI}} = \frac{1}{n}\begin{pmatrix} 0{,}824 & 0 \\ 0 & 0{,}381 \end{pmatrix} \text{ a}$$

$$\text{a } \mathbf{\Sigma}_{\text{GAS}} = \frac{1}{n}\begin{pmatrix} 0{,}549 & 0 \\ 0 & 0{,}235 \end{pmatrix}. \qquad (31)$$

We can see from equations (31) that

$$\hat{\eta}_3$$

and

$$\hat{\eta}_4$$

are asymptotically uncorrelated for each of the above-mentioned quick estimators. It is also obvious that we prefer Gastwirth estimator to median and trimean ones in terms of skewness and kurtosis estimation in the case of large samples from an (almost) normal distribution.

2.5 Application to Lognormal Distribution

LQ estimators for the three-parameter lognormal distribution behave similarly to L-moment estimators. We get the following expressions for LQ-moments of the above distribution from equations (7)–(9) and (13)

$$\xi_1 = \theta + \exp(\mu)\,\tau_{p,\alpha}(X_{1:1}), \qquad (32)$$

$$\xi_2 = \frac{1}{2}\exp(\mu)\,[\tau_{p,\alpha}(X_{2:2}) - \tau_{p,\alpha}(X_{1:2})], \qquad (33)$$

The LQ-skewness coefficient can be calculated using

$$\eta_3 = \frac{\frac{1}{3}\left[\tau_{p,\alpha}(X_{3:3}) - 2\,\tau_{p,\alpha}(X_{2:3}) + \tau_{p,\alpha}(X_{1:3})\right]}{\frac{1}{2}\left[\tau_{p,\alpha}(X_{2:2}) - \tau_{p,\alpha}(X_{1:2})\right]}. \qquad (34)$$

LQ-parameter estimators

$$\hat{\mu},\ \hat{\sigma} \text{ and } \hat{\theta}$$

represent the solution of equations (7)–(9) in combination with equations (32)–(34) for μ, σ and θ, where we replace ξ_r with $\hat{\xi}_r$.

Conducting regression analysis, we obtain the following approximate relationship, allowing for estimation of

$$\hat{\sigma}$$

for $|\eta_3| \leq 1{,}0$ a $|k| \leq 2{,}64$

$$\hat{\sigma} = 2{,}1684\hat{\eta}_3 + 0{,}3967\,\hat{\eta}_3^3 + 0{,}1744\hat{\eta}_3^5 - 0{,}1015\hat{\eta}_3^7. \qquad (35)$$

Once we get the value

$$\hat{\sigma},$$

we can also obtain estimates

$$\hat{\mu} \text{ and } \hat{\theta}$$

using equations (33) and (32).

2.6 Appropriateness of the Model

It is also necessary to assess the suitability of the constructed model or choose another model from several alternatives, applying a criterion which can be the sum of absolute deviations of the observed and theoretical frequencies for all intervals

$$S = \sum_{i=1}^{k} |n_i - n\,\pi_i| \qquad (36)$$

or criterion χ^2

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - n\,\pi_i)^2}{n\,\pi_i}, \qquad (37)$$

where n_i are the observed frequencies in individual intervals, π_i are the theoretical probabilities of statistical unit membership in the i-th interval, n is the total sample size of the corresponding statistical file, $n \cdot \pi_i$ are theoretical frequencies in individual intervals, $i = 1, 2, ..., k$, and k is the number of intervals.

The appropriateness of the wage distribution curve is not a general mathematical-statistical issue for testing the null hypothesis

H_0: the sample coming from the supposed theoretical distribution,

against the alternative hypothesis

H_1: non H_0,

because in wage distribution goodness of fit testing we often work with large samples, tests usually leading to the null hypothesis rejection. This results not only from the lower test power at a given significance level but also from the test construction itself. The smallest distribution deviations revealed by the test being of no practical significance, the general consistency between the model and reality proves sufficient for us to "borrow" the wage distribution curve, allowing for the tentative use of the χ^2 test criterion. Relying on experience and logical insights, however, the model suitability assessment remains to a large extent subjective.

3 Results and Discussion

The research database – consisting of employees who were working in the Czech Republic over the period 2009–2016 – is broken down by various demographic and socio-economic factors. The research variable is the gross monthly (nominal) wage (in CZK). Data were drawn from the official website of the Czech Statistical Office. They are in the form of the interval frequency distribution (a total of 328 wage distributions) since the data on individual employees are not available.

Tables 1–3 present parameter estimates obtained using the three methods of point parameter estimation and the *S*-criterion. Generally, the method of LQ-moments yielded the best results, deviations occurring mainly at both ends of the wage distribution due to extreme open intervals. With respect to total wage distribution sets, LQ-moments always give the most accurate outcomes in terms of the *S*-criterion. In the research of all 328 wage distributions, the method of TL-moments produced the second most accurate results in more than half of the cases, deviations occurring again especially at both ends of the distribution. The above tables indicate that TL-moments brought the second most accurate results in terms of all total sets of wage distributions for the Czech Republic in the period 2009–2016, the method of L-moments yielding the third most accurate outcomes in most cases.

Table 1: Parameter estimates obtained using LQ-moments and *S*-criterion values for total wage distribution in the Czech Republic

Year	Parameter estimation		
	μ	σ^2	θ
2009	9.059 747	0.630 754	9,065.52
2010	9.215 324	0.581 251	8,552.10
2011	9.277 248	0.573 002	8,872.54
2012	9.313 803	0.577 726	9,382.66
2013	9.382 135	0.680 571	10,027.84
2014	9.438 936	0.688 668	10,898.39
2015	9.444 217	0.703 536	10,640.53
2016	9.482 060	0.681 258	10,616.80

Year	*S*-criterion
2009	108,437.01
2010	146,509.34
2011	137,422.05
2012	149,144.68
2013	198,670.74
2014	206,698.93
2015	193,559.55
2016	202,367.04

Source: Own research

Table 2: Parameter estimates obtained using TL-moments and *S*-criterion values for total wage distribution in the Czech Republic

Year	Parameter estimation		
	μ	σ^2	θ
2009	9.017 534	0.608 369	7,664.46
2010	9.241 235	0.507 676	6,541.16
2011	9.283 399	0.515 290	6,977.45
2012	9.283 883	0.543 225	7,868.21
2013	9.387 739	0.601 135	7,902.64
2014	9.423 053	0.624 340	8,754.64
2015	9.431 478	0.631 013	8,684.51
2016	9.453 027	0.621 057	8,746.20

Year	*S*-criterion
2009	133,320.79
2010	248,438.78
2011	231,978.79
2012	216,373.24
2013	366,202.87
2014	357,668.48
2015	335,999.20
2016	323,851,84

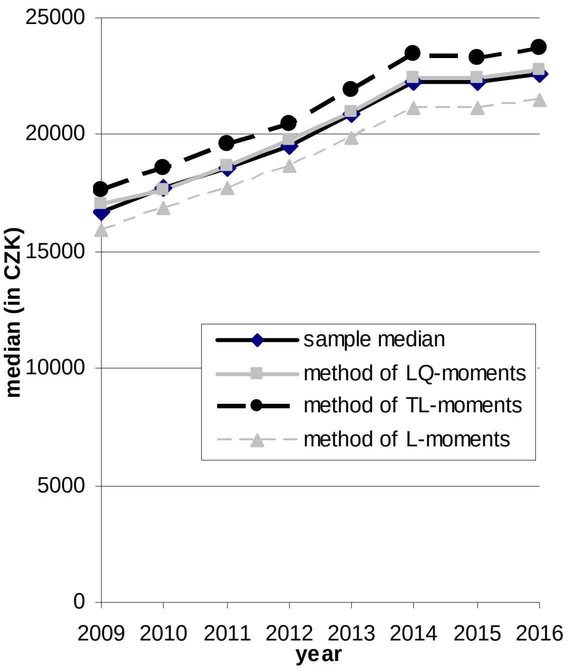
Source: Own research

Table 3: Parameter estimates obtained using L-moments and *S*-criterion values for total wage distribution in the Czech Republic

Year	Parameter estimation		
	μ	σ^2	θ
2009	9.741 305	0.197 395	2.07
2010	9.780 008	0.232 406	0.22
2011	9.833 604	0.228 654	0.27
2012	9.890 594	0.210 672	0.59
2013	9.950 263	0.268 224	0.16
2014	10.017 433	0.264 124	0.19
2015	10.019 787	0.269 047	0.20
2016	10.033 810	0.269 895	0.20
Year	***S*-criterion**		
2009	248,331.74		
2010	281,541.41		
2011	311,008.23		
2012	325,055.67		
2013	370,373.62		
2014	391,346.02		
2015	359,736.37		
2016	389,542.21		

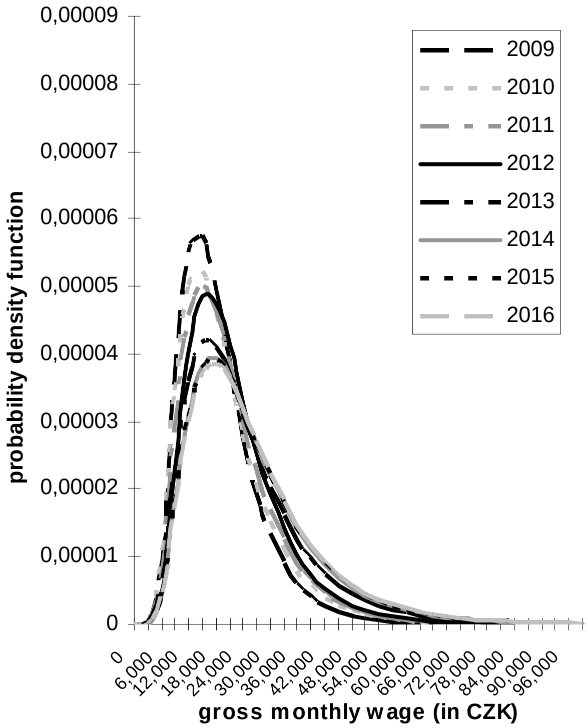
Source: Own research

Figure 1: Development of sample and theoretical median of three-parameter lognormal curves with parameters estimated using different methods of estimation



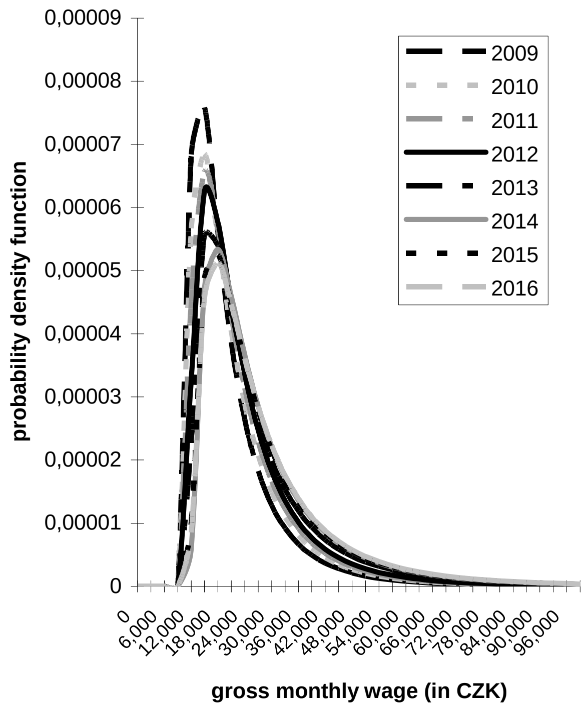
Source: Own research

Figure 2: Development of probability density function of three-parameter lognormal curves with parameters estimated using LQ-moments



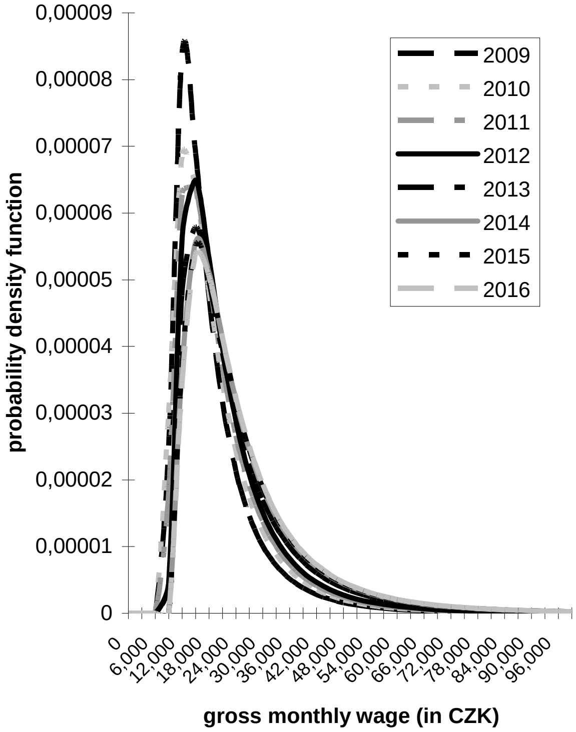
Source: Own research

Figure 3: Development of probability density function of three-parameter lognormal curves with parameters estimated using TL-moments



Source: Own research

Figure 4: Development of probability density function of three-parameter lognormal curves with parameters estimated using L-moments



Source: Own research

Figure 1 also gives some idea of the accuracy of point parameter estimation methods, showing the development of the sample median of gross monthly wage for the total set of all employees in the Czech Republic over the period 2009–2016 and the corresponding theoretical median of model three-parameter lognormal curves with parameters estimated by three various methods. The figure indicates that the curve showing the development of the theoretical median of the three-parameter lognormal distribution with parameters estimated using LQ-moments adheres the closest to that showing the course of the sample median. The other two curves tracing the development of the theoretical median of three-parameter lognormal curves with parameters estimated by the methods of TL- and L-moments are relatively distant from the course of the sample wage distribution median.

Figures 2–4 illustrate the development of the probability density function of three-parameter lognormal curves with parameters estimated by all three methods of LQ-, TL- and L-moments, again representing the course of model distributions of total wages earned by all employees in the Czech Republic between 2009 and 2016. We can see that the shapes of lognormal curves with parameters estimated using TL- and L-moments are similar to each other, while being markedly different from those whose parameters were estimated by the method of LQ-moments (cf. Figs. 3, 4 and 2, respectively).

4 Conclusion

The present paper deals with an alternative moment analysis of probability distributions. The method of LQ-moments is compared with those of L- and TL-moments, particularly in terms of their parameter estimation accuracy, using the sum of all absolute deviations of the observed and theoretical frequencies for all intervals as a criterion. The higher accuracy of LQ-moments approach was proved by examining the set of 328 wage distributions, advantages of TL-moments compared to L-moments being also confirmed. The values of the χ^2 criterion having been calculated for each wage distribution, the test always led to the rejection of the zero hypothesis about the supposed shape of the wage distribution because of the typically large sample sizes.

References:

[1] Abu El-Magd, N. A. T., TL-Moments of the Exponentiated Generalized Extreme Value Distribution, *Journal of Advanced research*, Vol. 1, No. 4, 2010, pp. 351–359.

[2] Ashour, S. K., El-Sheik, A. A., & Abu El-Magd, N. A. T., TL-Moments and LQ-Moments of the Exponentiated Pareto Distribution, *Journal of Scientific Research & Reports*, Vol. 4, No. 4, 2015, pp. 328–347.

[3] Bhuyan, A., & Borah, M., LQ-Moments for Regional Flood Frequancy Analysis: A Case Study for the North-Bank Region of the

Brahmaputra River, India, *Journal of Modern Applied Statistical Methods*, Vol. 10, No. 2, 2011, pp. 730–740.

[4] David, H. A., & Nagaraja, H. N., *Order Statistics*, John Wiley & Sons, 2003.

[5] Deka, S., Borah, M., & Kakaty, S. C., Distributions of Annual Maximum Rainfall Series of North-East India, *European Water*, Vol. 27, No. 28, 2009, pp. 3–14.

[6] Kandeel, A. F., Linear Moments: An Overview, *International Journal of Business and Statistical Analysis*, Vol. 2. No. 2, 2015, pp. 85–90.

[7] Mudholkar, G. S., & Hutson, A., LQ-Moments: Analogs of L-moments, *Journal of Statistical Planning and Inference*, Vol. 71, No. 1, 1998, pp. 191–208.

[8] Mudholkar, G. S., & Natarajan, R., Measures for Distributional Classification and Model Selection, in Balakrishnan, N., *Advances on Methodological and Applied Aspects of Probability and Statistics*, Taylor & Francis, 2002, pp. 87–100.

[9] Shabri, A., & Jemain, A. A., LQ-Moments: Application to the Extreme Value Type I Distribution, *Journal of Applied Sciences*, Vol. 6, No. 5, 2006a, pp. 993–997.

[10] Shabri, A., & Jemain, A. A., LQ-Moments: Application to the Log-Normal Distribution, *Journal of Mathematics and Statistics*, Vol. 2, No. 3, 2006b, pp. 1549–1560.

[11] Shabri, A., & Jemain, A. A., LQ-Moments: Parameter Estimation for Kappa Distribution, *Sains Malaysiana*, Vol. 39, No. 5, 2010, pp. 845–850.

[12] Šimková, T., & Picek, J., A Comparison of L-, LQ-, TL-Moment and Maximum Likelihood High Quantile Estimates of the GPD and GEV Distribution, *Communications in Statistics – Simulation and Computation*, Vol. 46, No. 8, 2017, pp. 5991–6010.

[13] Zaher, H. M., El-Sheik, A. A., & Abu El-Magd, N. A. T., Estimation of Pareto Parameters Using a Fuzzy Least-Squares Method and Other Known Techniques with a Comparison, *British Journal of Mathematics & Computer Science*, Vol. 4, No. 14, 2014, pp. 2067–2088.

[14] Zin, W. Z. W., & Jemain, A. A., Non-Parametric Quantile Selection for Extreme Distributions, *Journal of Modern Applied Statistical Methods*, Vol. 7, No. 2, 2008, pp. 454–466.

[15] Zin, W. Z. W., Jemain, A. A., & Ibrahim, K., The Best Fitting Distribution of Annual Maximum Rainfall in Peninsular Malaysia Based on Methods of L-Moment and LQ-Moment, *Theoretical and Applied Climatology*, Vol. 96, No. 3–4, 2009, pp. 337–344.

Acknowledgement:

This paper was subsidized by the funds of institutional support of a long-term conceptual advancement of science and research (no. IP400040) at the Faculty of Informatics and Statistics, University of Economics, Prague, Czech Republic.