The quantile-based generalized logistic distribution

Brenda V. Omachar*
Department of Statistics, University of Pretoria, Pretoria, South Africa – Brenda.Omachar@up.ac.za

Paul J. van Staden
Department of Statistics, University of Pretoria, Pretoria, South Africa – paul.vanstaden@up.ac.za

Abstract

This paper proposes a new quantile-based generalized logistic distribution. This four-parameter distribution is highly flexible with respect to distributional shape in that it explains extensive levels of skewness and kurtosis through the inclusion of two shape parameters. The distribution is characterized through its L-moments and an estimation algorithm is presented for estimating the distribution's parameters with method of L-moments estimation.

Keywords: L-moments; quantile function.

1. Introduction

A quantile-based distribution is specified in terms of its quantile function, $Q(p)$. This is because there exist no closed-form expressions for its cumulative distribution function, $F(x)$, or its probability density function, $f(x)$. Well-known examples of quantile-based distributions include Tukey's lambda distribution (Tukey, 1960), various types of generalized lambda distributions (Ramberg *et al.*, 1979; Freimer *et al.*, 1988; van Staden, 2013) and the Davies distribution (Hankin and Lee, 2006).

In this paper the quantile function of the generalized logistic distribution (GLO), proposed by Hosking (1986), is used as the building block to create the quantile function of a quantile-based GLO, henceforth denoted GLO_{QB}. This new four-parameter distribution is defined in Section 2 and its probabilistic properties are discussed. In Section 3 the GLO_{QB} is characterized through its L-moments. Section 4 presents an estimation algorithm using method of L-moments, which is then used in Section 5 to fit the GLO_{QB} to a data set.

2. Definition and properties

Following Hosking (1986), the quantile function of the standard GLO is given by

$$Q(p) = \frac{1}{\lambda}\left(1 - \left(\frac{1-p}{p}\right)^{\lambda}\right), \quad \lambda \neq 0. \qquad (1)$$

As shown by Gilchrist (2000), the distribution with quantile function $-Q(1-p)$ is the reflection of the distribution with quantile function $Q(p)$. Therefore, the quantile function of the reflected standard GLO is

$$Q(p) = \frac{1}{\lambda}\left(\left(\frac{p}{1-p}\right)^{\lambda} - 1\right), \quad \lambda \neq 0. \qquad (2)$$

The properties of the GLO were studied by Hosking (1986). Specifically he presented expressions for the L-moments and the conventional moments of the GLO, where the expressions of the L-moments are considerably simpler than those of the conventional moments.

<u>**Definition 1.**</u> *Let X be a real-valued random variable. X is said to have a quantile-based generalized logistic distribution, denoted* $X \sim GLO_{QB}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$, *if it has the quantile function*

$$Q(p) = \lambda_1 + \frac{1}{\lambda_2}\left(\frac{1}{\lambda_3}\left(\left(\frac{p}{1-p}\right)^{\lambda_3} - 1\right) - \frac{1}{\lambda_4}\left(\left(\frac{1-p}{p}\right)^{\lambda_4} - 1\right)\right), \quad 0 < p < 1, \qquad (3)$$

where λ_1 *and* λ_2 *are respectively location and scale parameters and* λ_3 *and* λ_4 *are shape parameters.*

The probabilistic properties of the GLO_{QB} are briefly discussed below.

- The probability density curve of the GLO_{QB} is J-shaped when $\lambda_3 \geq 1$ and $\lambda_4 \leq -1$ and also when $\lambda_3 \leq -1$ and $\lambda_4 \geq 1$, while it is unimodal for all other combinations of values of λ_3 and λ_4.
- When $\lambda_3 = \lambda_4 = 0$, the GLO_{QB} reduces to the logistic distribution.
- The GLO_{QB} is symmetric for $\lambda_3 = \lambda_4$, positively skewed for $\lambda_3 > \lambda_4$ and negatively skewed for $\lambda_3 < \lambda_4$.
- For all values of λ_3 and λ_4, the GLO_{QB} is leptokurtic (heavy-tailed).
- The GLO_{QB} has infinite support, $(-\infty, \infty)$, for $\lambda_3 \leq 0$ and $\lambda_4 \leq 0$ as well as for $\lambda_3 \geq 0$ and $\lambda_4 \geq 0$. When $\lambda_3 < 0$ and $\lambda_4 > 0$, the support of the GLO_{QB} is $\left(-\infty, \lambda_1 + \frac{1}{\lambda_2}\left(\frac{1}{\lambda_4} - \frac{1}{\lambda_3}\right)\right)$, while it is $\left(\lambda_1 + \frac{1}{\lambda_2}\left(\frac{1}{\lambda_4} - \frac{1}{\lambda_3}\right), \infty\right)$ if $\lambda_3 > 0$ and $\lambda_4 < 0$. In effect, the GLO_{QB} has half-infinite support when the signs of the shape parameters differ.
- The GLO_{QB} with parameters $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ is the reflection of the GLO_{QB} with parameters $(\lambda_1, \lambda_2, \lambda_4, \lambda_3)$ about the line $x = \lambda_1$.
- The GLO_{QB} with parameters $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ is the same as the GLO_{QB} with parameters $(\lambda_1, \lambda_2, -\lambda_4, -\lambda_3)$.

3. *L*-moments

L-moments, defined by Hosking (1990), are expectations of linear combinations of order statistics. The first order *L*-moment, L_1, is the *L*-location, while L_2 is the *L*-scale. The r^{th} order *L*-moment ratio is given by $\tau_r = \frac{L_r}{L_2}$, with τ_3 and τ_4 the *L*-skewness and *L*-kurtosis ratios respectively.

Theorem 1. *If $X \sim GLO_{QB}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ with $-1 < \lambda_3 < 1$ and $-1 < \lambda_4 < 1$, then*

$$L_1 = \lambda_1 + \frac{1}{\lambda_2}\left(\pi(\mathrm{Csc}[\pi\lambda_3] - \mathrm{Csc}[\pi\lambda_4]) - \frac{1}{\lambda_3} + \frac{1}{\lambda_4}\right), \quad (4)$$

$$L_2 = \frac{1}{\lambda_2}\left(\pi(\mathrm{Csc}[\pi\lambda_3]\lambda_3 + \mathrm{Csc}[\pi\lambda_4]\lambda_4)\right), \quad (5)$$

$$\tau_3 = \frac{\mathrm{Csc}[\pi\lambda_3]\lambda_3^2 - \mathrm{Csc}[\pi\lambda_4]\lambda_4^2}{\mathrm{Csc}[\pi\lambda_3]\lambda_3 + \mathrm{Csc}[\pi\lambda_4]\lambda_4}, \quad (6)$$

and

$$\tau_4 = \frac{\mathrm{Csc}[\pi\lambda_3]\lambda_3(1+5\lambda_3^2) + \mathrm{Csc}[\pi\lambda_4]\lambda_4(1+5\lambda_4^2)}{6(\mathrm{Csc}[\pi\lambda_3]\lambda_3 + \mathrm{Csc}[\pi\lambda_4]\lambda_4)}, \quad (7)$$

where $\mathrm{Csc}[\cdot]$ *is the cosec function.*

4. Method of *L*-moments estimation

Given an ordered data set of sample size *n*, $x_{1:n} \leq x_{2:n} \leq \ldots \leq x_{n:n}$, method of *L*-moments estimates, $(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\lambda}_4)$, for the four parameters of the GLO_{QB} can be obtained using the following estimation algorithm:

Step 1: Calculate the first four sample *L*-moments using

$$\ell_r = \binom{n}{r}^{-1} \sum\sum\cdots\sum_{1\leq i_1<i_2<\ldots<i_r\leq n} r^{-1}\sum_{k=0}^{r-1}(-1)^k\binom{r-1}{k} x_{i_{r-k}:n}, \quad (8)$$

for $r = 1,2,3,4$ (Hosking, 1990) and then the sample *L*-skewness and *L*-kurtosis ratios, $t_3 = \frac{\ell_3}{\ell_2}$ and $t_4 = \frac{\ell_4}{\ell_2}$. Verify whether the values of t_3 and t_4 lie within the (τ_3, τ_4) space of the GLO_{QB}. If so, proceed with Step 2. If not, the GLO_{QB} cannot be fitted to the data.

Step 2: Since both the *L*-skewness and the *L*-kurtosis ratios of the GLO_{QB} are functions of both shape parameters, utilize a numerical optimization method such as the *FindRoot* function in *Mathematica* 10.0 (Wolfram, 2014) to simultaneously obtain solutions for $\hat{\lambda}_3$ and $\hat{\lambda}_4$ using Eqs. (6) and (7).

Step 3: Solve for $\hat{\lambda}_2$ using Eq. (5) and then for $\hat{\lambda}_1$ using Eq. (4).

5. Example

Fig. 1(a) shows a histogram for the concentrations of polychlorinated biphenyl (PCB) in the yolk lipids of pelican eggs, used in Thas (2010) as an example data set with respect to goodness-of-fit testing. For this data set of $n = 65$ observations, the values of the sample L-location, L-scale, L-skewness ratio and L-kurtosis ratio are $\ell_1 = 210$, $\ell_2 = 39.793$, $t_3 = 0.104$ and $t_4 = 0.213$ respectively. The method of L-moments estimates for the fitted GLO_{QB} are $\hat{\lambda}_1$=203.598, $\hat{\lambda}_2$=0.055, $\hat{\lambda}_3$=0.301 and $\hat{\lambda}_4$=0.12. The probability density curve of the fitted GLO_{QB}, plotted in Fig. 1(a), as well as the Q-Q plot in Fig. 1(b) indicate that the GLO_{QB} provides an adequate fit to the data set.

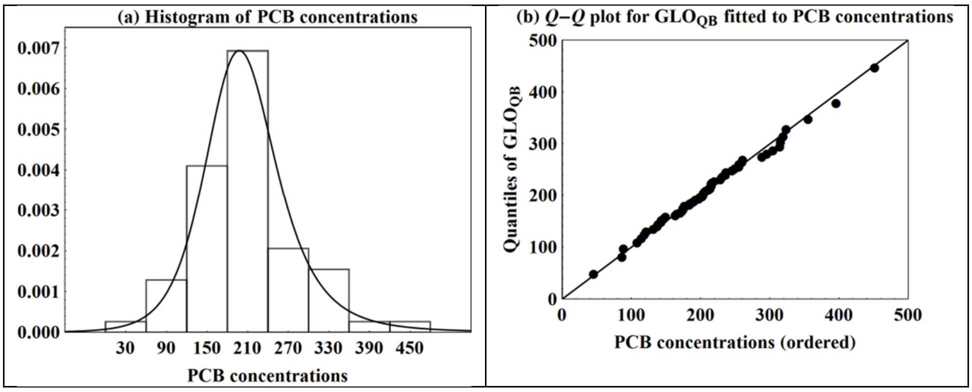

Fig. 1. A histogram of the PCB concentrations of polychlorinated biphenyl (PCB) in the yolk lipids of pelican eggs together with the probability density curve of the fitted GLO_{QB} and the corresponding Q-Q plot.

References

Freimer, M., Mudholkar, G.S., Kollia, G., & Lin, C.T. (1988). A study of the generalized Tukey lambda family, *Communications in Statistics: Theory and Methods*, 17(10), 3547–3567.

Gilchrist, W.G. (2000). *Statistical Modelling with Quantile Functions*, Chapman & Hall / CRC Press, Boca Raton, Florida.

Hankin, R.K.S., & Lee, A. (2006). A new family of non-negative distributions, *Australian & New Zealand Journal of Statistics*, 48(1), 67–78.

Hosking, J.R.M. (1986). The theory of probability weighted moments, *Research Report RC12210*, *IBM Research*, *Yorktown Heights*.

Hosking, J.R.M. (1990). *L*-moments: Analysis and estimation of distributions using linear combinations of order statistics, *Journal of the Royal Statistical Society, Series B (Methodological)*, 52(1), 105–124.

Ramberg, J.S., Tadikamalla, P.R., Dudewicz, E.J., & Mykytka, E.F. (1979). A probability distribution and its uses in fitting data, *Technometrics*, 21(2), 201–214.

Thas, O. (2010). *Comparing Distributions*, Springer, New York.

Tukey, J.W. (1960). The practical relationship between the common transformations of percentages of counts and of amounts, *Technical Report 36*, *Statistical Techniques Research Group*, *Princeton University*.

van Staden, P.J. (2013). Modeling of generalized families of probability distributions in the quantile statistical universe, *PhD Thesis*, *Department of Statistics*, *University of Pretoria*.

Wolfram. (2014). *Mathematica*, *Version 10.0*. Wolfram Research, Inc., Champaign, Illinois.