University of Victoria
Department of Economics

Econometrics Working Paper EWP0901

ISSN 1485-6441

Bias Reduction for the Maximum Likelihood Estimator of the Scale Parameter in the Half-Logistic Distribution

David E. Giles

Department of Economics, University of Victoria
Victoria, B.C., Canada V8W 2Y2

Revised, May 2009

Abstract

We derive an analytic expression for the bias, to $O(n^{-1})$ of the maximum likelihood estimator of the scale parameter in the half-logistic distribution. Using this expression to bias-correct the estimator is shown to be very effective in terms of bias reduction, without adverse consequences for the estimator's precision. The analytic bias-corrected estimator is also shown to be dramatically superior to the alternative of bootstrap-bias-correction.

Keywords Half-logistic distribution; Life testing; Bias reduction

Mathematics Subject Classification 62F10; 62F40; 62N02; 62N05

Author Contact:
David E. Giles, Dept. of Economics, University of Victoria, P.O. Box 1700, STN CSC, Victoria, B.C., Canada V8W 2Y2; e-mail: dgiles@uvic.ca; Phone: (250) 721-8540; FAX: (250) 721-6214

1. Introduction

The half-logistic (or folded-logistic) distribution was proposed by Balakrishnan (1985) as a life-testing model. One of the attractions of this distribution in the context of reliability theory is that it has a monotonically increasing hazard rate for all parameter values, a property shared by relatively few distributions which have support on the positive real half-line. In terms of tail behaviour, the half-logistic distribution provides a degree of flexibility as its tail thickness lies between those of the half-normal and half-Cauchy distributions. The half-logistic distribution has also been used successfully to model records. For example, Mbah and Tsokos (2008) apply it to environmental and sports records data.

If X follows a logistic distribution, then $Y = |X|$ has a half-logistic distribution, the p.d.f. for which is:

$$\begin{aligned} f(y) &= \frac{(2/\sigma)\exp\{-(y-\mu)/\sigma\}}{[1+\exp\{-(y-\mu)/\sigma\}]^2} \\ &= \frac{(2/\sigma)\exp\{(y-\mu)/\sigma\}}{[1+\exp\{(y-\mu)/\sigma\}]^2} \end{aligned} \quad ; \qquad y \ge \mu > 0\,, \;\; \sigma > 0, \qquad (1)$$

where μ and σ are the location and scale parameters respectively. The moments of this distribution are given in Appendix A.

Various estimators for the parameters of the half-logistic distribution have been proposed, for both uncensored and censored data. For example, see Balakrishnan and Puthenpura (1986), Balakrishnan and Wong (1991), Balakrishnan and Chan (1992), and Adatia (1997, 2000). In addition, the operating characteristic under acceptance sampling from the half-logistic distribution has been discussed by Kantam and Rosaiah (1998).

In this paper we deal with maximum likelihood estimation with uncensored data. If the location parameter of (1) is unknown, its MLE is $\hat{\mu} = Y_{1:n}$, where $Y_{j:n}$ is the j^{th} order statistic in a sample of size n (Balakrishnan and Wong, 1991, p.142). However, the MLE for the scale parameter cannot be expressed in closed form. Notwithstanding this complication, we derive the bias, to $O(n^{-1})$, of the MLE of the scale parameter of the half-logistic distribution in the interesting case where the location parameter is zero. It transpires that the MLE has extremely small relative bias, even in very small samples. We also consider a simple bias-corrected counterpart to this estimator, and show that its bias is an order of magnitude less than that of the MLE itself, and that this is achieved without loss of precision. The alternative of using the bootstrap to reduce the bias of the MLE is found to be totally inferior to our analytic correction.

The next section presents summarizes the techniques used to evaluate the bias to $O(n^{-1})$. Our principal results appear in section 3; and some numerical evaluations are given in section 4. Section 5 concludes. Some technical details are provided in Appendix A.

2. Preliminary Results

Let $l(\theta)$ be the log-likelihood function based on a sample of *n* observations, with *p*-dimensional parameter vector, θ. $l(\theta)$ is assumed to be regular with respect to all derivatives up to and including the third order. Define:

$$k_{ij} = E(\partial^2 l / \partial\theta_i \partial\theta_j) \quad ; \quad i, j = 1, 2, \ldots, p \tag{2}$$

$$k_{ijl} = E(\partial^3 l / \partial\theta_i \partial\theta_j \partial\theta_l) \quad ; \quad i, j, l = 1, 2, \ldots, p \tag{3}$$

$$k_{ij,l} = E[(\partial^2 l / \partial\theta_i \partial\theta_j)(\partial l / \partial\theta_l)] \quad ; \quad i, j, l = 1, 2, \ldots, p\ . \tag{4}$$

and

$$k_{ij}^{(l)} = \partial k_{ij} / \partial\theta_l \quad ; \quad i, j, l = 1, 2, \ldots, p. \tag{5}$$

All of the expressions in (2) – (5) are assumed to be $O(n)$. Extending earlier work by Tukey (1949), Bartlett (1953a, 1953b), Haldane (1953), Haldane and Smith (1956), Shenton and Wallington (1962) and Shenton and Bowman (1963), Cox and Snell (1968) showed that when the sample data are independent (but not necessarily identically distributed) the bias of the s^{th} element of the MLE of θ ($\hat{\theta}$) is:

$$Bias(\hat{\theta}_s) = \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{l=1}^{p} k^{si} k^{jl} [0.5 k_{ijl} + k_{ij,l}] + O(n^{-2}) \ ; \quad s = 1, 2, \ldots, p. \tag{6}$$

where k^{ij} is the $(i,j)^{\text{th}}$ element of the inverse of the (expected) information matrix, $K = \{-k_{ij}\}$. Cordeiro and Klein (1994) note that this bias expression also holds if the data are non-independent, provided that all of the *k* terms are $O(n)$, and show that it can be re-written as:

$$Bias(\hat{\theta}_s) = \sum_{i=1}^{p} k^{si} \sum_{j=1}^{p} \sum_{l=1}^{p} [k_{ij}^{(l)} - 0.5 k_{ijl}] k^{jl} + O(n^{-2}) \ ; \ s = 1, 2, \ldots, p. \tag{7}$$

The computational advantage of (7) is that it does not involve terms of the form defined in (4).

Now, let $a_{ij}^{(l)} = k_{ij}^{(l)} - (k_{ijl} / 2)$, for $i, j, l = 1, 2, \ldots, p$; and define the following matrices:

$$A^{(l)} = \{a_{ij}^{l}\} \ ; \quad i, j, l = 1, 2, \ldots, p \tag{8}$$

$$A = [A^{(1)} \,|\, A^{(2)} \,|\, ...\,|\, A^{(p)}]\,. \tag{9}$$

Cordeiro and Klein (1994) show that the expression for the $O(n^{-1})$ bias of $\hat{\theta}$ can be re-written as:

$$Bias(\hat{\theta}) = K^{-1} A \; vec(K^{-1}) + O(n^{-2})\;. \tag{10}$$

A "bias-corrected" MLE for θ can then be obtained as:

$$\tilde{\theta} = \hat{\theta} - \hat{K}^{-1} \hat{A} \; vec(\hat{K}^{-1})\,, \tag{11}$$

where $\hat{K} = (K)|_{\hat{\theta}}$ and $\hat{A} = (A)|_{\hat{\theta}}$. It can be shown that the bias of $\tilde{\theta}$ will be $O(n^{-2})$. It is crucial to note that (10) and (11) can be evaluated even when the likelihood equation does not admit a closed-form analytic solution, so that the MLE has to be obtained *via* a numerical solution.

3. Bias of the MLE for the Shape Parameter

Under independent sampling from the half-logistic distribution, with uncensored data, the log-likelihood function (when $\mu = 0$) is:

$$l = n\ln(2) - n\ln(\sigma) + (n\bar{y} / \sigma) - 2\sum_{i=1}^{n} \ln[1 + \exp(y_i / \sigma)]\,, \tag{12}$$

where $\bar{y} = n^{-1} \sum_{i=1}^{n} y_i$. So,

$$\partial l / \partial \sigma = -(n / \sigma) - (n\bar{y} / \sigma^2) + (2 / \sigma^2) \sum_{i=1}^{n} [y_i \exp(y_i / \sigma)] / [1 + \exp(y_i / \sigma)] \quad . \tag{13}$$

Note that there is no closed-form solution to the likelihood equation obtained by equating (13) to zero.

In what follows, we will require the following higher-order derivatives of the log-likelihood function:

$$\begin{aligned} \partial^2 l / \partial \sigma^2 = (n / \sigma^2) + (2n\bar{y} / \sigma^3) - (4 / \sigma^3) \sum_{i=1}^{n} [y_i \exp(y_i / \sigma)] / [1 + \exp(y_i / \sigma)] \\ - (2 / \sigma^4) \sum_{i=1}^{n} [y_i^2 \exp(y_i / \sigma)] / [1 + \exp(y_i / \sigma)]^2 \end{aligned} \tag{14}$$

$$\partial^3 l/\partial\sigma^3 = -(2n/\sigma^3) - (6n\bar{y}/\sigma^4) + (12/\sigma^4)\sum_{i=1}^{n}[y_i \exp(y_i/\sigma)]/[1+\exp(y_i/\sigma)]$$
$$+ (12/\sigma^5)\sum_{i=1i}^{n}[y_i^2 \exp(y_i/\sigma)]/[1+\exp(y_i/\sigma)]^2 \qquad (15)$$
$$+ (2/\sigma^6)\sum_{i=1i}^{n}[y_i^3\{\exp(y_i/\sigma) - \exp(2y_i/\sigma)\}]/[1+\exp(y_i/\sigma)]^3$$

To evaluate the expectations of these derivatives we will use the following results for a half-logistic variate, *Y*, the proofs of which appear in Appendix B:

$$E\{[y\exp(y/\sigma)]/[1+\exp(y/\sigma)]\} = \sigma[\ln(2)+0.5] \qquad (16)$$

$$E\{[y^2\exp(y/\sigma)]/[1+\exp(y/\sigma)]^2\} = (\sigma^2/3)[(\pi^2/6)-1] \qquad (17)$$

$$E\{[y^3(\exp(y/\sigma)-\exp(2y/\sigma))]/[1+\exp(y/\sigma)]^3\} = \sigma^3[0.5-(\pi^2/12)] \qquad (18)$$

We then have the following expressions relating to the joint cumulants of the derivatives of the log-likelihood function:

$$\begin{aligned} k_{11} &= E[\partial^2 l/\partial\sigma^2] = -(n/9\sigma^2)(3+\pi^2) \\ &= -1.429956044(n/\sigma^2) \end{aligned} \qquad (19)$$

$$\begin{aligned} k_{111} &= E[\partial^3 l/\partial\sigma^3] = (n/2\sigma^3)(2+\pi^2) \\ &= 5.934802199(n/\sigma^3) \end{aligned} \qquad (20)$$

In addition,

$$\begin{aligned} k_{11}^{(1)} &= \partial k_{11}/\partial\sigma = 2n(3+\pi^2)/(9\sigma^3) \\ &= 2.859912088(n/\sigma^3) \end{aligned} \qquad (21)$$

and

$$\begin{aligned} a_{11}^{(1)} &= k_{11}^{(1)} - 0.5k_{111} \\ &= -0.107489011(n/\sigma^3) \end{aligned} \qquad (22)$$

The (expected) information measure is

$$K = -k_{11} = 1.429956044(n/\sigma^2) \quad , \qquad (23)$$

and

$$A = a_{11}^{(1)} \quad . \qquad (24)$$

So, using Cordeiro and Klein's (1994) modification of the Cox-Snell (1968) result, to $O(n^{-1})$,

$$Bias(\hat{\sigma}) = K^{-1}Avec(K^{-1}) \\ = -0.052567665(\sigma / n) \tag{25}$$

The bias is unambiguously negative, and small in relative terms. Moreover, the relative bias is invariant to the value of *σ*. An unbiased (to $O(n^{-2})$) estimator of *σ* is $\tilde{\sigma} = (\hat{\sigma} - Bias(\hat{\sigma})) = \hat{\sigma}(n + 0.052567665)/n$. Of course, correcting for the bias in this way also has implications for the mean squared error (MSE) of the estimator, and this point is taken up in the next section.

4. Numerical Evaluations

The bias expression in (25) is valid only to $O(n^{-1})$. The actual bias and mean squared error (MSE) of the maximum likelihood and bias-corrected maximum likelihood estimators have been simulated in a Monte Carlo experiment. The simulations were undertaken using the *maxLik* package (Toomet, 2008) for the *R* statistical software environment (R, 2008). Half-logistic variates were generated using the inversion method, and the log-likelihood function was maximized using the Nelder-Mead algorithm.

In addition to $\hat{\sigma}$ and $\tilde{\sigma}$, we have also considered the bootstrap-bias-corrected estimator (Efron, 1979). The latter is obtained as $\breve{\sigma} = 2\hat{\sigma} - (1/N_B)[\sum_{j=1}^{N_B} \hat{\sigma}_{(j)}]$, where $\hat{\sigma}_{(j)}$ is the MLE of σ obtained from the j^{th} of the N_B bootstrap samples. See Efron (1982, p.33). This estimator is also unbiased to $O(n^{-2})$, but in many applications it is known to suffer from excessive variance.

Each part of the experiment relating to $\hat{\sigma}$ and $\tilde{\sigma}$ uses 250,000 Monte Carlo replications. In the case of $\breve{\sigma}$ the number of Monte Carlo replications is 250,000, with 1,000 bootstrap samples *per* replication (*i.e*., 250 million evaluations for each value of *n*, in this case). The results that are reported in Table 1 are percentage biases and MSE’s, the latter being defined as $100 \times$(MSE / σ^2). For each of the estimators under consideration, both of these measures are invariant to the value of σ, for a given sample size.

Several key results emerge from Table 1. First, the percentage bias of the MLE is negative but small, even for very small sample sizes, which is encouraging for users of this estimator. Second, however, the absolute bias of the bias-corrected estimator, $\tilde{\sigma}$, is often an order of magnitude less than that of the MLE. This bias-corrected estimator is trivial to implement, and provides dramatic gains - its bias is negligible, in percentage terms. Third, these gains in bias reduction when using $\tilde{\sigma}$ come at the cost of increases in variance, as is evidenced by the very small differences in the percentage MSE's that are reported for $\hat{\sigma}$ and $\tilde{\sigma}$. Fourth, the bootstrap-bias-corrected estimator performs essentially as well as $\tilde{\sigma}$, in terms of bias reduction (especially when nthe sample size exceeds 50) but at some computational expense. Finally, however, using the bootstrap correction is slightly less effective than is the analytic correction in terms of mean squared error. Overall, these results strongly favour using the Cox-Snell analytic approach to correct for bias to $O(n^{-1})$.

5. Conclusion

The maximum likelihood estimator of the scale parameter in the half-logistic distribution has a very small (negative) percentage bias, even with quite small sample sizes. However, using the Cox-Snell procedure for determining the $O(n^{-1})$ bias of this estimator, and then making the associated analytic bias correction, proves to be highly effective. The percentage bias is often reduced by an order of magnitude, with essentially no cost in terms of increased mean squared error. In terms of bias reduction, it is generally at least as effective as using the bootstrap to bias-correct the maximum likelihood estimator, and usually slightly better than the bootstrap correction in terms of mean squared error. The analytic bias correction is extremely simple to apply in practice, and is recommended.

Acknowledgment

The author is grateful to Helen Feng for helpful discussions and suggestions; Ryan Godwin for programming assistance; and Lief Bluck for arranging access to the required computing facilities.

Table 1

Simulated percentage biases and mean squared errors for the maximum likelihood and bias-corrected maximum likelihood estimators

n	$\%\,Bias(\hat{\sigma})$	$\%\,Bias(\tilde{\sigma})$	$\%\,Bias(\breve{\sigma})$	$\%\,MSE(\hat{\sigma})$	$\%\,MSE(\tilde{\sigma})$	$\%\,MSE(\breve{\sigma})$
10	-0.4827	0.0404	0.0988	6.9512	7.0221	7.0402
15	-0.3279	0.0214	-0.0390	4.6267	4.6581	4.6793
20	-0.2400	0.0223	0.0415	3.4784	3.4961	3.5016
25	-0.1719	0.0380	0.0331	2.7966	2.8081	2.7997
30	-0.1370	0.0380	0.0166	2.3271	2.3351	2.3342
35	-0.1206	0.0294	0.0141	1.9962	2.0020	2.0010
40	-0.1098	0.0264	0.0077	1.7505	1.7549	1.7540
45	-0.0902	0.0239	-0.0259	1.5549	1.5584	1.5607
50	-0.0811	0.0214	0.0135	1.3996	1.4025	1.4022
75	-0.0497	0.0203	0.0014	0.9300	0.9313	0.9334
100	-0.0337	0.0188	-0.0073	0.6988	0.6995	0.7008
150	-0.0170	0.0180	-0.0137	0.4651	0.4655	0.4678
200	-0.0137	0.0126	-0.0093	0.3502	0.3504	0.3498
250	-0.0133	0.0077	0.0040	0.2808	0.2809	0.2806

Appendix A

A.1 *Moments of the half-logistic distribution*

The central moments of the standardized half-logistic distribution can be evaluated using the following result (Gradshteyn, and Ryzhik, 1994; integral 3.424, no.2):

$$\int_0^{\infty} \frac{(1+a)e^x + a}{(1+e^x)^2} e^{-ax} x^n dx = n! \sum_{k=0}^{\infty} \frac{(-1)^k}{(a+k+1)^n} \quad . \tag{A.1}$$

Letting $Z = (Y - \mu)/\sigma$, from (1) the density of the standardized variate, *Z*, is

$$f(z) = \frac{2e^z}{[1+e^z]^2} \qquad ; \qquad z > 0$$

and

$$E(Z^r) = 2\int_0^{\infty} \frac{x^r e^z}{[1+e^z]^2} dz \quad .$$

Applying (A.1) with $a = 0$ and $n = r$:

$$E(Z^r) = 2r! \sum_{k=0}^{\infty} (-1)^k (k+1)^{-r} \ . \tag{A.2}$$

An alternative derivation of (A.2) is provided by Balakrishnan and Wong (1991, p. 140). Using the Maclaurin series for $\ln(1+w)$ with $w = 1$, it follows that $E(Z) = \ln(4)$. Similarly, using the relationship $(\pi^2/6) = \sum_{k=1}^{\infty} k^{-2}$, we have $E(Z^2) = (\pi^2/3)$. The moments of *Y* itself can, of course, be derived directly from (A.2) by applying the binomial theorem:

$$E(Y^r) = 2r!\sigma^r \sum_{j=0}^{r} {}^rC_j (\mu/\sigma)^{r-j} \sum_{k=0}^{\infty} (-1)^k (k+1)^{-j} \ . \tag{A.3}$$

When $\mu = 0$, using the convention that $0^0 = 1$, it follows immediately that $E(Y) = 2\sigma \ln(2)$ and $V(Y) = \sigma^2\{(\pi^2/3) - 4[\ln(2)]^2\}$, *etc*.

A.2 *Derivation of equations (16) – (18)*

All of the following results have been established analytically, and then verified by using the Maple 10 package (Maplesoft, 2005). Further details are available on request.

(i) Equation (16) follows directly from (13) by recalling that $E(\partial l / \partial \sigma) = 0$, and using the result that $E(Y) = 2\sigma \ln(2)$, from Appendix A.

(ii) $E\{[y\exp(y/\sigma)]/[1+\exp(y/\sigma)]\} = (2/\sigma)\int_0^\infty \frac{y^2 \exp(2y/\sigma)}{[1+\exp(y/\sigma)]^4} dy$.

The evaluation of this integral is tedious, but can be accomplished by using the change of variable, $z = \exp(y/\sigma)$, and then repeatedly integrating by parts and by partial fractions. A final change of variable, $w = z^{-1}$, and the use of the integral $\int_0^1 \frac{\ln(1+w)}{w} dw = (\pi^2/12)$ (Gradshteyn, and Ryzhik, 1994; integral 4.291, no.1.) , yields the result:

$E\{[y\exp(y/\sigma)]/[1+\exp(y/\sigma)]\} = (\sigma^2/3)[(\pi^2/6)-1]$.

(iii) $E\{[y^3(\exp(y/\sigma)-\exp(2y/\sigma))]/[1+\exp(y/\sigma)]^3\}$

$$= (2/\sigma)\int_0^\infty \frac{y^3[\exp(2y/\sigma)-\exp(3y/\sigma)]}{[1+\exp(y/\sigma)]^5} dy .$$

This integral can be evaluated by using the same approach as in (ii), yielding: $E\{[y^3(\exp(y/\sigma)-\exp(2y/\sigma))]/[1+\exp(y/\sigma)]^3\} = \sigma^3[0.5-(\pi^2/12)]$.

References

Adatia, A. (1997). Approximate BLUE's of the parameters of the half logistic distribution based on fairly large doubly censored samples. *Computational Statistics and Data Analysis* 24: 179-191.

Adatia, A. (2000). Estimation of the parameters of the half-logistic distribution using generalized rank set sampling. *Computational Statistics and Data Analysis* 33: 1-13.

Balakrishnan, N. (1985). Order statistics from the half logistic distribution. *Journal of Statistical Computation and Simulation* 20: 287-309.

Balakrishnan, N., Asgharzadeh, A. (2005). Inference for the scaled half-logistic distribution based on progressively type-II censored samples. *Communications in Statistics – Theory and Methods* 34: 73-87.

Balakrishnan, N., Chan, P. S. (1992). Estimation for the scaled half logistic distribution under Type II censoring. *Computational Statistics and Data Analysis* 13: 123-131.

Balakrishnan, N., Puthenpura, S. (1986). Best linear unbiased estimators of location and scale parameters of the half logistic distribution. *Journal Statistical Computation and Simulation* 25: 193–204.

Balakrishnan, N., Wong, K. H. T. (1991). Approximate MLEs for the location & scale parameters of the half-logistic distribution with type-II censoring. *IEEE Transactions on Reliability* 40: 140-145.

Bartlett, M. S. (1953a). Approximate confidence intervals. *Biometrika* 40: 12-19.

Bartlett, M. S. (1953b). Approximate confidence intervals II. More than one unknown parameter. *Biometrika* 40: 306-317.

Cordeiro, G. M., Klein, R. (1994). Bias correction in ARMA models. *Statistics and Probability Letters* 19: 169-176.

Cox, D. R., Snell, E. J. (1968). A general definition of residuals. *Journal of the Royal Statistical Society, B* 30: 248-275.

Cribari-Neto, F., Vasconcellos, K. L. P. (2002). Nearly unbiased maximum likelihood estimation for the beta distribution. *Journal of Statistical Computation and Simulation* 72: 107-118.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics* 7: 1-26.

Efron, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Gradshteyn, I. S., Ryzhik, I. W. (1994). *Table of Integrals, Series, and Products* (ed. A. Jeffrey), 5th ed. New York: Academic Press.

Haldane, J. B. S. (1953). The estimation of two parameters from a sample. *Sankhyā* 12: 313-320.

Haldane, J. B. S., Smith, S. M. (1956). The sampling distribution of a maximum likelihood estimate. *Biometrika* 43: 96-103.

Kantam, R. R. L., Rosaiah, K. (1998). Half logistic distribution in acceptance sampling based on life tests. *IAPQR Transactions* 23: 117–125.

Maplesoft (2005), *Maple 10*. Waterloo, ON: Waterloo Maple.

Mbah, A. K., Tsokos, C. P. (2008). Record values from half logistics and inverse Weibull probability distribution functions. *Neural, Parallel and Scientific Computations* 16: 73-92.

R (2008). The R Project for Statistical Computing, http://www.r-project.org

Shenton, L. R., Bowman, K. (1963). Higher moments of a maximum-likelihood estimate. *Journal of the Royal Statistical Society, B* 25: 305-317.

Shenton, L. R., Wallington, P. A. (1962). The bias of moment estimators with an application to the negative binomial distribution. *Biometrika* 49: 193-204.

Toomet, O., Henningsen, A. (2008). *maxLik: Maximum likelihood estimation*. http://CRAN.R-project.org; http://maxLik.org

Tukey, J. W. (1949). Standard confidence points. Memorandum Report No. 26, Statistical Research Group, Stanford University.