Improved Maximum Likelihood Estimation for the Akash Distribution

David E. Giles

Department of Economics

University of Victoria, CANADA

November, 2024

Abstract:

We consider three analytical methods for reducing the finite-sample bias of the maximum likelihood estimator of the (scale) parameter in the Akash distribution. The latter distribution has flexible features that make it attractive for modelling lifetime data. Based on a simulation experiment, all three bias-reduction methods are found to be highly effective, and have the added merit of also reducing the mean squared error of the maximum likelihood estimator. The analytical results are also illustrated with six real-life data-sets.

Key Words: Akash distribution; lifetime data; maximum likelihood estimation; bias reduction

Author contact:

58 Rock Lake Court, RR1, Bancroft, Ontario K0L 1C0, CANADA; dgiles@uvic.ca ; +1-613-202 9501

(*The author is Professor Emeritus, Department of Economics, University of Victoria, CANADA*.)

1. Introduction

Among the many statistical distributions that are used in reliability studies to analyze lifetime data, the Akash distribution proposed by Shanker (2015) has certain advantages. It is simple, with just one (scale) parameter, but its density and hazard functions have more flexible shapes than those for competing distributions, such as the exponential and Lindley distributions. Notwithstanding its simplicity, maximum likelihood estimation of the scale parameter involves numerical optimization, and the resulting estimator is biased in finite samples.

In this paper we consider three analytical methods for reducing the order of magnitude of this bias. Specifically, we compare the "corrective" method of Cox and Snell (1968), Firth's (1993) "preventive" method, and a more recent corrective method proposed by Godwin and Giles (2019) that allows for less restrictive bias functions. These modified maximum likelihood estimators (MLE's) are compared in a Monte Carlo simulation experiment. The results show that as well as reducing bias, all three estimators also reduce the (percentage) mean squared error (MSE) of the original MLE. Of the three estimators, the Godwin-Giles estimator performs best in terms of bias reduction, but the other two estimators have a slight advantage in terms of MSE.

The Akash distribution and the analytical results relating to the various estimators are introduced in the next section. Section 3 outlines the simulation experiment and its associated results, and some illustrative applications with real data are provided in section 4. Some concluding remarks appear in section 5.

2. Theoretical results

If X follows the Akash distribution with a scale parameter, λ, its density function is

$$f(x) = \frac{\lambda^3}{(\lambda^2+2)}(1+x^2)exp(-\lambda x) \qquad ; \qquad x > 0, \lambda > 0 \tag{1}$$

and its distribution function is

$$F(x) = 1 - \left[1 + \frac{\lambda x(\lambda x+2)}{(\lambda^2+2)}\right] \exp(-\lambda x) \; . \tag{2}$$

Shanker (2015, p.67) shows that the r'th. moment of X about the origin is

$$\mu_r' = r!\,[\lambda^2 + (r+1)(r+2)]/[\lambda^r(\lambda^2+2)] \qquad ; \quad r = 1, 2, 3, \ldots. \tag{3}$$

implying that $E[X] = [\lambda^2 + 6]/[\lambda(\lambda^2+2)]$; and $Var.[X] = [\lambda^4 + 16\lambda^2 + 12]/[\lambda(\lambda^2+2)]^2$.

This distribution has been generalized in various ways by authors such as Shanker and Shukla (2017), Shanker *et al*. (2018), and Abushal (2021). However, the basic Ashal distribution has considerable flexibility and is the focus of this paper.

Using a sample of n independent observations, with a sample mean of $\bar{x}$, the log-likelihood function based on (1) is

$$l = 3nlog(\lambda) - nlog(\lambda^2 + 2) - \lambda n\bar{x} + \sum_{i=1}^{n} \log(1 + x_i^2). \quad (4)$$

So,

$$\frac{\partial l}{\partial \lambda} = \left(\frac{3n}{\lambda}\right) - \frac{2n\lambda}{(\lambda^2+2)} - n\bar{x}\ , \quad (5)$$

and the MLE ($\tilde{\lambda}$) of λ is obtained by (numerically) solving the equation,

$$\bar{x}\lambda^3 - \lambda^2 + 2\bar{x}\lambda - 6 = 0 \quad (6)$$

for λ. Equating $\bar{x}$ and the above expression for $E[X]$, it follows immediately that the MLE of λ is also the method of moments estimator for that parameter.

The MLE for λ is consistent, asymptotically unbiased, and asymptotically efficient. However, its finite-sample properties have not been explored previously. In this paper we obtain an analytical approximation to the small-sample bias of this estimator, and derive three (approximately) "bias-corrected" estimators. The biases and MSEs of the latter estimators are compared with the corresponding properties of $\tilde{\lambda}$ itself in an extensive simulation experiment, described in the next section.

In what follows, we require the following results:

$$\frac{\partial^2 l}{\partial \lambda^2} = -\left(\frac{3n}{\lambda^2}\right) + \frac{4n\lambda^2}{(\lambda^2+2)^2} - \frac{2n}{(\lambda^2+2)} \quad (7)$$

$$\frac{\partial^3 l}{\partial \lambda^3} = \left(\frac{6n}{\lambda^3}\right) - \frac{16n\lambda^3}{(\lambda^2+2)^3} + \frac{12n\lambda}{(\lambda^2+2)^2} \quad (8)$$

and we define:

$$k_{11} = E\left[\frac{\partial^2 l}{\partial \lambda^2}\right] = \frac{\partial^2 l}{\partial \lambda^2} \quad (9)$$

$$k_{111} = E\left[\frac{\partial^3 l}{\partial \lambda^3}\right] = \frac{\partial^3 l}{\partial \lambda^3} \quad (10)$$

and

$$k_{11}^{(1)} = \frac{k_{11}}{\partial\lambda} = \frac{\partial^3 l}{\partial\lambda^3} \quad . \tag{11}$$

The information measure is given by

$$K = -k_{11} = \left(\frac{3n}{\lambda^2}\right) - \frac{4n\lambda^2}{(\lambda^2+2)^2} + \frac{2n}{(\lambda^2+2)} \tag{12}$$

and so

$$k^{11} = K^{-1} = [\lambda(\lambda^2+2)]^2/[3n(\lambda^2+2)^2 - 4n\lambda^4 + 2n\lambda^2(\lambda^2+2)]. \tag{13}$$

If we define

$$A = a_{11} = k_{11}^{(1)} - 0.5k_{111} = 0.5k_{111} = \left(\frac{3n}{\lambda^3}\right) - \frac{8n\lambda^3}{(\lambda^2+2)^3} + \frac{6n\lambda}{(\lambda^2+2)^2} \ , \tag{14}$$

Then, following Cordeiro and Klein (1994), the bias of the MLE, $\tilde{\lambda}$, can be expressed as

$$B(\tilde{\lambda}) = \frac{A}{K^2} = \frac{n^{-1}\lambda(\lambda^2+2)[3(\lambda^2+2)^3 - 8\lambda^6 + 6\lambda^4(\lambda^2+2)]}{[3(\lambda^2+2)^2 - 4\lambda^4 + 2\lambda^2(\lambda^2+2)]^2} + O(n^{-2}) \quad . \tag{15}$$

The Cox-Snell bias-corrected estimator of λ is

$$\hat{\lambda} = \tilde{\lambda} - \tilde{B}(\tilde{\lambda}) \quad , \tag{16}$$

where $\tilde{B}(\tilde{\lambda}) = B(\tilde{\lambda})_{(\lambda=\tilde{\lambda})}$.

Firth's bias-corrected estimator, $\check{\lambda}$, is obtained by solving the equation,

$$\frac{\partial l}{\partial\lambda} - K(\lambda)B(\lambda) = 0 \tag{17}$$

for λ.

The Cox-Snell and Firth estimators implicitly make the strong assumption that the bias function is "flat". Godwin and Giles (2019) allow for a bias correction that avoids this restrictive assumption by proposing the estimator

$$\ddot{\lambda} = \tilde{\lambda} - \ddot{B}(\tilde{\lambda}), \tag{18}$$

where $\ddot{B}(\tilde{\lambda}) = B(\tilde{\lambda})|_{\lambda=\ddot{\lambda}}$.

In (18), the bias function is evaluated at $\ddot{\lambda}$ (which is unbiased), rather than evaluated at $\tilde{\lambda}$ (which is biased). Of course, the difficulty with $\ddot{\lambda}$ is that there is no closed-form solution for this expression. Giles and Godwin rearrange (18), to get

$$\tilde{\lambda} = \ddot{\lambda} + \ddot{B}(\tilde{\lambda}), \tag{19}$$

and then the estimator $\ddot{\lambda}$ may then be obtained by substituting $[\lambda + B(\tilde{\lambda})]$ for λ in the log-likelihood function, (4), and maximizing the latter numerically.

3. Simulation experiment

A Monte Carlo simulation experiment has been conducted to evaluate and compare the percentage biases and percentage MSEs of the original MLE ($\tilde{\lambda}$) and the three bias-corrected estimators ($\hat{\lambda}, \check{\lambda}$ and $\ddot{\lambda}$), for sample sizes ranging from $n = 10$ to $n = 100$, and values of λ between 0.05 and 5.0. These parameter values allow for a range of shapes for the Akash density, and reflect the estimated values obtained in a number of applications. For example, see Shanker (2015) and Shanker and Fesshave (2016). All of the simulations used *NREP* = 50,000 replications and were carried out using the R programming language (R Core Team, 2024). Random variates for the Akash distribution were generated by the acceptance-rejection method, using the R package 'AcceptReject' (Marinho, 2024). The author's R code (including that associated with the applications in section 4) can be downloaded from https://github.com/DaveGiles1949/r-code.

The simulated percentage biases and percentage MSEs of $\tilde{\lambda}$ are calculated as

$$\%Bias(\tilde{\lambda}) = 100\left[\left(\frac{1}{NREP}\sum_{j=1}^{NREP}\tilde{\lambda}_j\right) - \lambda\right]/\lambda$$

and

$$\%MSE(\tilde{\lambda}) = 100\left[\frac{1}{NREP}\sum_{j=1}^{NREP}\left(\tilde{\lambda}_j - \lambda\right)^2\right]/\lambda^2$$

respectively, where $\tilde{\lambda}_j$ is the *j*'th. replication of the estimator, $\tilde{\lambda}$. The same calculations are made for the estimators, $\hat{\lambda}, \check{\lambda}$ and $\ddot{\lambda}$.

The simulation results are reported in Table 1. There, we see that the (unadjusted) MLE of the scale parameter is biased upwards. In percentage terms, this bias (and the percentage MSE) increases with the true value of that parameter, for any given sample size. In all cases, the percentage biases and MSE's decrease as *n* increases, reflecting the consistency of all of the estimators. The Cox-Snell and Firth estimators reduce the (absolute) percentage bias of the original MLE by a similar substantial degree, and often by an order of magnitude when *n* is small. Often, this results in a small negative bias in estimation. In all cases, the reduction in bias is achieved with the added benefit of a decrease in the percentage MSE. The Godwin-Giles estimator generally has the smallest absolute percentage bias. It also has smaller percentage MSE than the original MLE, but generally its percentage MSE is slightly greater than those of the other two bias-corrected estimators, $\hat{\lambda}$ and $\check{\lambda}$.

4. Empirical applications

We present six empirical applications to illustrate the consequences of applying the different bias-corrected MLEs of the Akash distribution's scale parameter. The summary statistics for the data used in these applications, named "Data 1" to "Data 6" in Table 2, are reported in that table and can be downloaded from https://github.com/DaveGiles1949/Data.

Also shown in Table 2 are the values of the Anderson-Darling (A-D) statistic for testing the goodness-of-fit of the Akash distribution to each data-set. Here, the A-D test is one that has been modified using the "bias transformation" approach proposed by Rasche (2009) for the beta and gamma distributions, and introduced by Giles (2024) for the Akash distribution. The latter author shows that this A-D test out-performs other (modified) goodness-of-fit tests based on the empirical distribution function. The results in Table 2 suggest that the hypothesis that "Data 1" follows the Akash distribution should be rejected, at the 5% significance level; but this distribution is supported for the other samples at this level of significance.

"Data 1" is the sample used in the first application in Shanker (2015), originally reported by Gross and Clark (1975, p.105). The data measure relief times (in minutes) of 20 patients treated with an analgesic medication. "Data 2" is Shanker's second application data-set, and it measures the breaking strength of aircraft window glass. The sample of 31 observations is from Fuller *et al*. (1994). The "Data 3" sample comprises 34 observations on the amount of vinyl chloride found in clean upgradient monitoring wells, measured in mg/litre This data set was reported by Bhaumik *et al*. (2009), and is Dataset 11 in the applications of the Akash distribution reported by Shanker and Fesshave (2016). "Data 4" was also used by Bhaumik *et al*. (2009, Table 4), and measures the survival time (in weeks) of 20 male mice that were exposed to gamma radiation. "Data 5" and "Data 6" are taken from the 'reliability' data set in the R package, 'survival' (Therneau, 2024). The first of these samples comprises 11 observations on the time of inspection for turbine wheels for cracks, in hundreds of hours. The second sample (with n =70) measures the time-to-failure of diesel generator fans (in hours).

The basic MLE ($\tilde{\lambda}$) and the three "bias-corrected" MLEs for the scale parameter of the Akash distribution are also reported in Table for each application. In all cases, the bias-corrected estimates are essentially the same in value, and generally slightly smaller in value than the basic MLE. However, viewed in percentage terms, the differences between the value of $\tilde{\lambda}$ and the bias-corrected estimates are approximately -1.7%, -1.0%, -1.0%, -1.5%, and -3.0% for the first five application. In the case of "Data 6", this difference is between -1.2% and +3.0%, depending on the choice of bias correction

5. Conclusions

In this paper we have considered the finite-sample properties of the maximum likelihood estimator of the Akash distribution's sole parameter. This estimator is positively biased, and various ways of eliminating the first-order bias have been considered. The results of an extensive simulation experiment show that the Cox and Snell (968) "corrective" approach, and Firth's (1993) "preventive" approach produce similar, and very successful, results in terms of reduced percentage bias and percentage mean squared error. The Godwin-Giles (2019) approach, that allows for a less restrictive bias function, performs somewhat better than the other two approaches in terms of bias reduction, but at the expense of slightly less improvement in percentage mean squared error. Overall, the application of one or other of the bias corrections is recommended.

Several applications involving real-life data illustrate the extent to which each "bias correction" alters the numerical values of the estimates of the Akash distribution's scale parameter in relatively small samples. Decreases in the estimates' values by an order of 1% to 3% is found to be typical.

Table 1: %Bias and %MSE of estimators

n	%Bias ($\tilde{\lambda}$) [%MSE ($\tilde{\lambda}$)]	%Bias ($\hat{\lambda}$) [%MSE ($\hat{\lambda}$)]	%Bias ($\check{\lambda}$) [%MSE ($\check{\lambda}$)]	%Bias ($\ddot{\lambda}$) [%MSE ($\ddot{\lambda}$)]
		λ = 0.5		
10	2.3869	-0.7314	-0.7321	-0.6438
	[3.0247]	[2.8020]	[2.8032]	[2.8067]
15	1.5750	-0.4861	-0.4859	-0.4459
	[1.9291]	[1.8342]	[1.8345]	[1.8354]
25	1.0335	-0.1963	-0.1955	-0.1864
	[1.1380]	[1.1021]	[1.1021]	[1.1024]
50	0.4499	-0.1614	-0.1604	-0.1590
	[0.5439]	[0.5360]	[0.5360]	[0.5361]
100	0.2066	-0.0983	-0.0973	-0.0977
	[0.2662]	[0.2643]	[0.2643]	[0.2643]
		λ = 1.0		
10	2.9443	-0.5122	-0.4727	-0.3565
	[3.3418]	[2.9412]	[2.9504]	[2.9649]
15	1.7167	-0.5306	-0.5147	-0.4638
	[2.0442]	[1.8917]	[1.8941]	[1.8977]
25	1.1456	-0.1843	-0.1796	-0.1609
	[1.1793]	[1.1238]	[1.1244]	[1.1251]
50	0.5275	-0.1290	-0.1290	-0.1218
	[0.5641]	[0.5514]	[0.5515]	[0.5515]
100	0.2581	-0.0683	-0.0696	-0.0663
	[0.2726]	[0.2696]	[0.2695]	[0.2696]
		λ = 1.5		
10	3.9390	-0.5226	-0.4255	-0.2159
	[4.1763]	[3.4194]	[3.4370]	[3.4800]
15	2.5449	-0.3291	-0.2877	-0.1972
	[2.4733]	[2.1745]	[2.1793]	[2.1906]
25	1.4393	-0.2397	-0.2252	-0.1934
	[1.3695]	[1.2723]	[1.2732]	[1.2754]
50	0.6071	-0.2157	-0.2120	-0.2044
	[0.6347]	[0.6138]	[0.6138]	[0.6141]
100	0.3230	-0.0853	-0.0842	-0.0811
	[0.3103]	[0.3050]	[0.3050]	[0.3051]

Table 1 (continued): %Bias and %MSE of estimators

n	%Bias ($\tilde{\lambda}$) [%MSE ($\tilde{\lambda}$)]	%Bias ($\hat{\lambda}$) [%MSE ($\hat{\lambda}$)]	%Bias ($\check{\lambda}$) [%MSE ($\check{\lambda}$)]	%Bias ($\ddot{\lambda}$) [%MSE ($\ddot{\lambda}$)]
		$\lambda = 2.0$		
10	5.2479	-0.4400	-0.2981	0.0537
	[5.5470]	[4.2711]	[4.2909]	[4.3817]
15	3.2108	-0.4164	-0.3548	-0.2033
	[3.1275]	[2.6476]	[2.6534]	[2.6764]
25	1.8957	-0.2152	-0.1932	-0.1399
	[1.6645]	[1.5082]	[1.5095]	[1.5140]
50	0.8653	-0.1654	-0.1595	-0.1468
	[0.7622]	[0.7273]	[0.7274]	[0.7279]
100	0.4714	-0.0390	-0.0368	-0.0354
	[0.3709]	[0.3619]	[0.3620]	[0.3620]
		$\lambda = 2.5$		
10	6.4486	-0.3886	-0.2294	0.2769
	[7.0442]	[5.2362]	[5.2469]	[5.3883]
15	4.0407	-0.3162	-0.2455	-0.0249
	[3.8170]	[3.1380]	[3.1419]	[3.1782]
25	2.3295	-0.1978	-0.1725	-0.0941
	[2.0100]	[1.7897]	[1.7907]	[1.7979]
50	1.1063	-0.1262	-0.1198	-0.1004
	[0.9101]	[0.8598]	[0.8599]	[0.8608]
100	0.5684	-0.0408	-0.0392	-0.0353
	[0.4382]	[0.4257]	[0.4257]	[0.4258]
		$\lambda = 3.0$		
10	7.4025	-0.3880	-0.2373	0.4096
	[8.3419]	[6.0948]	[6.0948]	[6.2748]
15	4.6785	-0.2938	-0.2259	0.0589
	[4.5845]	[3.7179]	[3.7179]	[3.7661]
25	2.7083	-0.1766	-0.1519	-0.0499
	[2.3293]	[2.0919]	[2.0521]	[2.0616]
50	1.2988	-0.1080	-0.1018	-0.0764
	[1.0494]	[0.9854]	[0.9855]	[0.9866]
100	0.6649	-0.0304	-0.0289	-0.0233
	[0.5035]	[0.4877]	[0.4877]	[0.4878]

Table 2: Maximum likelihood estimation results

	Data 1	Data 2	Data 3	Data 4	Data 5	Data 6
n	20	31	34	20	11	70
mean	1.90	30.81	1.88	113.45	25.82	4921.00
median	1.70	29.90	1.15	119.00	26.00	4300.00
s.d.	0.70	7.25	1.95	35.79	13.58	2837.41
A-D	0.8045^{**}	0.4509	0.6388^{*}	0.4297	0.2889	0.7383^{*}
$\tilde{\lambda}$	1.1569	0.0970	1.1657	0.0264	0.1157	$0.5919\text{x}10^{-3}$
$\hat{\lambda}$	1.1369	0.0960	1.1538	0.0260	0.1122	$0.5891\text{x}10^{-3}$
$\check{\lambda}$	1.1371	0.0960	1.1539	0.0260	0.1122	$0.6098\text{x}10^{-3}$
$\ddot{\lambda}$	1.1374	0.0960	1.1540	0.0260	0.1123	$0.5845\text{x}10^{-3}$

Note: 10% and 5% A-D critical values are 0.631and 0.752. See Stephens (1986, p.123), Table 4.7.
*Significant at the 10% level; **significant at the 5% level.

References

Abushal, T. A., 2021. Parametric inference of Akash distribution for type-II censoring with analyzing of relief times of patients. *Mathematics*, 6, 10789–10801.

Bhaumik, D. K., K. Kapur, and R. D. Gibbons, 2009. Testing parameters of a gamma distribution for small samples. *Technometrics*, 51, 326- 334.

Cordeiro, G. M. and F. Cribari-Neto, 2014. *An Introduction to Bartlett Correction and Bias Reduction*. Springer: Heidelberg.

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

Cox, D. R. and E. J. Snell, 1968. A General definition of residuals. *Journal of the Royal Statistical Society, Series B*, 30, 248–275.

Firth, D., 1993. Bias reduction of maximum likelihood estimates. *Biometrika*, 80, 27-38.

Fuller, E. J., S. Frieman, J. Quinn, G. Quinn, and W. Carter, 1994. Fracture mechanics approach to the design of glass aircraft windows: A case study. *Society of Photo-Optical Instrumentation Engineers Proceedings*, 2286, 419-430.

Giles, D. E., 2024. Goodness-of-Fit Tests for the Akash Distribution. Unpublished working paper. https://github.com/DaveGiles1949/My-Documents/blob/master/GOF%20for%20Akash.pdf.

Godwin, R. T. and D. E. Giles, 2019. Improved analytic bias correction for maximum likelihood estimators. *Communications in Statistics – Simulation and Computation*, 48, 15-26.

Gross, A. J. and V. A. Clark, 1975. *Survival Distributions: Reliability Applications in the Biometrical Sciences*, Wiley, New York.

Marinho, P. R. D., 2024. R package ‘AcceptReject’: Acceptance-rejection method for generating pseudo-random observations. https://cran.r-project.org/web/packages/AcceptReject/index.html.

R Core Team, 2024. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Raschke, M., 2009. Biased transformation and its application in goodness-of-fit tests for the beta and gamma distribution. *Communications in Statistics - Simulation and Computation*, 38, 1870–1890.

Shanker, R., 2015. Akash distribution and its applications. *International Journal of Probability and Statistics*, 4, 65-75.

Shanker, R. and H. Fesshaye, 2016. On modeling of lifetime data using Akash, Shanker, Lindley and exponential distributions. *Biometrics and Biostatistics International Journal*, 3, 214-224.

Shanker, R. and K. K. Shukla, 2017. On two-parameter Akash distribution. *Biometrics and Biostatistics International Journal*, 6, 416-425.

Shanker, R., K. K. Shukla, R. Shanker, and A. Pratap, 2018. A generalized Akash distribution. *Biometrics and Biostatistics International Journal*, 7, 18-26.

Stephens, M. A., 1986. Tests based on EDF statistics. In: D'Augustino, R. B. and M. A. Stephens, eds., *Goodness-of-Fit Techniques*. Marcel Dekker, New York, 97-194.

Therneau, T. M., 2024. R package 'survival': Survival analysis. https://cran.r-project.org/web/packages/survival/index.html.