Journal of the Korean Data & Information Science Society
2009, **20**(6), 1199–1206

한국데이터정보과학회지

Estimation for the double Rayleigh distribution based on progressive Type-II censored samples

Suk-Bok Kang[1] · Won-Tae Jung[2]

[12]Department of Statistics, Yeungnam University

Received 15 September 2009, revised 14 November 2009, accepted 18 November 2009

Abstract

This paper deals with the estimation based on progressive Type-II censored samples from the double Rayleigh distribution. We derive some estimators of the location and scale parameters of the double Rayleigh distribution based on progressive Type-II censored samples. We compare the proposed estimators in the sense of the mean squared error for various censored samples.

Keywords: Approximate maximum likelihood estimator, double Rayleigh distribution, progressive Type-II censored sample.

1. Introduction

The probability density function (pdf) and the cumulative distribution function (cdf) of the random variable having the double Rayleigh distribution are given by

$$f(x) = \frac{|x-\theta|}{2\sigma^2}\exp\left[-\frac{(x-\theta)^2}{2\sigma^2}\right] \tag{1.1}$$

and

$$F(x) = \frac{1}{2}\left[1 + sgn(x)\left\{1 - \exp\left[-\frac{(x-\theta)^2}{2\sigma^2}\right]\right\}\right], \tag{1.2}$$

where σ and θ are the scale and the location parameters, respectively and

$$sgn(x) = \begin{cases} 1, & x \geq \theta, \\ -1, & x < \theta. \end{cases}$$

Dattatreya Rao and Narasimham (1989) obtained the best linear unbiased estimators (BLUEs) for the location and scale parameters with various shape parameter values in the

[1] Corresponding author: Professor, Department of Statistics, Yeungnam University, Gyeongsan 712-749, Korea. E-mail: sbkang@yu.ac.kr
[2] Master, Department of Statistics, Yeungnam University, Gyeongsan 712-749, Korea.

double Weibull distribution for complete and censored samples. Vasudeva *et al.* (1991) obtained the first two moments and product moments of the absolute value of the order statistics for the double exponential distribution and the double Weibull distribution. They also obtained the optimum unbiased absolute estimator of the scale parameter.

It has been noted that in most cases, the maximum likelihood method does not provide explicit estimators based on censored samples. Especially, when the sample is progressive censored, the maximum likelihood method does not admit explicit solutions. Hence it is desirable to develop approximations to this maximum likelihood method which would provide us with estimators that are explicit functions of order statistics.

The approximate maximum likelihood estimating method was first developed by Balakrishnan (1989) for the purpose of providing the explicit estimators of the scale parameter in the Rayleigh distribution. Kang *et al.* (2004) introduced the approximate maximum likelihood estimator (AMLE) of the scale parameter of the Weibull distribution based on multiply Type-II censored samples. Han and Kang (2006) derived the AMLEs of the scale parameter and the location parameter in the two-parameter Rayleigh distribution under multiply Type-II censoring by the approximate maximum likelihood estimation method when two parameters are unknown. Son and Woo (2007) defined a skew-symmetric double Rayleigh distribution and derived an AMLE and a moment estimator of a skewed parameter in a skew-symmetric double Rayleigh distribution. Kang (2007) proposed some explicit estimators for the half triangle distribution based on multiply Type-II censored samples. Recently, Lee *et al.* (2008) proposed the AMLEs of the scale parameter in a triangular distribution based on multiply Type-II censored samples by the approximate maximum likelihood estimation methods. Han and Kang (2008) derived the AMLEs of the scale and location parameters in a double Rayleigh distribution based on multiply Type-II censored samples. They also compared the proposed estimators in the sense of the mean squared error (MSE) for various censored samples.

In this paper, we derive the maximum likelihood estimator (MLE), the least square estimator (LSE), and the AMLEs of the location and scale parameters in the double Rayleigh distribution based on progressive Type-II censored samples. We also compare the proposed estimators in the sense of the MSE for various censored samples.

2. Maximum likelihood estimation

Let us consider the following progressive Type-II censoring scheme. Suppose n randomly selected units with double Rayleigh distribution in the equation (1.1) were placed on a life test, only m are completely observed until failure. At the time of the first failure, R_1 of the $n-1$ surviving units are randomly withdrawn (or censored) from the life-testing experiment. At the time of the next failures, R_2 of the $n-2-R_1$ surviving units are censored, and so on. Finally, at the time of the m -th failures, all the remaining $R_m = n-m-R_1-\cdots-R_{m-1}$ surviving units are censored. Let

$$X_{1:m:n} \leq X_{2:m:n} \leq \cdots \leq X_{m:m:n} \tag{2.1}$$

denote such a progressive Type-II censored sample with $(R_1, \ldots, R_m)$ being the progressive censoring scheme.The likelihood function based on progressive Type-II censored sample in the equation (2.1) is given by

$$L = C\prod_{i=1}^{m} f(x_{i:m:n};\sigma,\theta)[1-F(x_{i:m:n};\sigma,\theta)]^{R_i}, \tag{2.2}$$

where $C = n(n-1-R_1)(n-2-R_1-R_2)\cdots(n-m+1-R_1-\cdots-R_{m-1})$.

The random variable $Z = (X-\theta)/\sigma$ has a standard double Rayleigh distribution with pdf and cdf:

$$f(z) = \frac{|z|}{2}\exp\left(-\frac{z^2}{2}\right) \tag{2.3}$$

and

$$F(z) = \begin{cases} 1-\dfrac{1}{2}\exp\left(-\dfrac{z^2}{2}\right), & z \geq 0, \\ \dfrac{1}{2}\exp\left(-\dfrac{z^2}{2}\right), & z < 0. \end{cases} \tag{2.4}$$

The functions $f(z\ )$ and $f'(z)$ are satisfied as

$$\frac{f'(z)}{f(z)} = \frac{1}{z} - z.$$

From the equation (2.2), the log-likelihood function may be written as

$$\ln L = K - m\ln\sigma + \sum_{i=1}^{m}\ln f(z_{i:m:n}) + \sum_{i=1}^{m} R_i\ln[1-F(z_{i:m:n})], \tag{2.5}$$

where K is a constant.

On differentiating the log-likelihood function with respect to σ and θ in turn and the equations to zero, we obtain the estimating equations as

$$\frac{\partial\ln L}{\partial\sigma} = -\frac{1}{\sigma}\left[2m - \sum_{i=1}^{m} z_{i:m:n}^2 - \sum_{i=1}^{m} R_i\frac{f(z_{i:m:n})}{1-F(z_{i:m:n})}z_{i:m:n}\right] = 0 \tag{2.6}$$

and

$$\frac{\partial\ln L}{\partial\theta} = -\frac{1}{\sigma}\left[\sum_{i=1}^{m}\frac{1}{z_{i:m:n}} - \sum_{i=1}^{m} z_{i:m:n} - \sum_{i=1}^{m} R_i\frac{f(z_{i:m:n})}{1-F(z_{i:m:n})}\right] = 0. \tag{2.7}$$

We can find the MLE of σ as values $\widehat{\sigma}$ that maximize the log-likelihood function in (2.5) by solving the equation $\partial\ln L/\partial\sigma = 0$ and the MLE of θ as values $\widehat{\theta}$ that maximize the log-likelihood function in (2.5) by solving the equation $\partial\ln L/\partial\theta = 0$. Since the equations (2.6) and (2.7) can not be solved explicitly, a numerical method may be used for the numerical determination of the MLEs. We consider sample sizes of 20(10)40 and we evaluate the mean squared errors of the MLEs $\widehat{\sigma}$ and $\widehat{\theta}$ using the Newton-Rapson method. These values are given in Table 5.1 and 5.2.

3. Approximate maximum likelihood estimation

Since the log-likelihood equations are very complicated, the equations (2.6) and (2.7) do not admit explicit solutions for σ and θ. So we need some approximate likelihood equations which can be given explicit solutions.

Let

$$\xi_{i:m:n} = F^{-1}(p_{i:m:n}) = \begin{cases} \sqrt{-2\ln(2q_{i:m:n})}, & p_{i:m:n} \geq 0.5, \\ -\sqrt{-2\ln(2p_{i:m:n})}, & p_{i:m:n} < 0.5, \end{cases} \tag{3.1}$$

where $q_{i:m:n} = 1 - p_{i:m:n}$ and

$$p_{i:m:n} = 1 - \prod_{j=m-i+1}^{m} \frac{j + R_{m-j+1} + \cdots + R_m}{j + 1 + R_{m-j+1} + \cdots + R_m}, \quad i = 1, \ldots, m.$$

We can approximate the functions by Taylor series expansion as follows;

$$\frac{f(z_{i:m:n})}{1 - F(z_{i:m:n})} \simeq \alpha_{1i} + \beta_{1i} z_{i:m:n}, \tag{3.2}$$

$$\frac{f(z_{i:m:n})}{1 - F(z_{i:m:n})} z_{i:m:n} \simeq \alpha_{2i} + \beta_{2i} z_{i:m:n} \tag{3.3}$$

and

$$\frac{1}{z_{i:m:n}} \simeq \gamma_{1i} + \delta_{1i} z_{i:m:n}, \tag{3.4}$$

where

$$\alpha_{1i} = \begin{cases} 0, & p_{i:m:n} \geq 0.5, \\ -\dfrac{p_{i:m:n}}{q_{i:m:n}} \xi^3_{i:m:n} \left[1 + \dfrac{p_{i:m:n}}{q_{i:m:n}}\right], & p_{i::m:n} < 0.5, \end{cases}$$

$$\beta_{1i} = \begin{cases} 1, & p_{i:m:n} \geq 0.5, \\ -\dfrac{p_{i:m:n}}{q_{i:m:n}} \left[1 - \xi^2_{i:m:n} - \dfrac{p_{i:m:n}}{q_{i:m:n}} \xi^2_{i:m:n}\right], & p_{i:m:n} < 0.5, \end{cases}$$

$$\alpha_{2i} = \begin{cases} -\xi^2_{i:m:n}, & p_{i:m:n} \geq 0.5, \\ \dfrac{p_{i:m:n}}{q_{i:m:n}} \xi^2_{i:m:n} \left[1 - \xi^2_{i:m:n} - \dfrac{p_{i:m:n}}{q_{i:m:n}} \xi^2_{i:m:n}\right], & p_{i:m:n} < 0.5, \end{cases}$$

$$\beta_{2i} = \begin{cases} 2\xi_{i:m:n}, & p_{i:m:n} \geq 0.5, \\ -\dfrac{p_{i:m:n}}{q_{i:m:n}} \xi_{i:m:n} \left[2 - \xi^2_{i:m:n} - \dfrac{p_{i:m:n}}{q^2_{i:m:n}} \xi^2_{i:m:n}\right], & p_{i:m:n} < 0.5, \end{cases}$$

$$\gamma_{1i} = \frac{2}{\xi_{i:m:n}}, \quad \delta_{1i} = -\frac{1}{\xi^2_{i:m:n}}.$$

By substituting the equations (3.2), (3.3), and (3.4) into the equations (2.6) and (2.7), we may approximate the equations (2.6) and (2.7) by

$$\frac{\partial \ln L}{\partial \sigma} \approx -\frac{1}{\sigma}\left[2m - \sum_{i=1}^{m} z_{i:m:n}^2 - \sum_{i=1}^{m} R_i(\alpha_{2i} + \beta_{2i} z_{i:m:n})\right] = 0 \tag{3.5}$$

and

$$\frac{\partial \ln L}{\partial \theta} \approx -\frac{1}{\sigma}\left[\sum_{i=1}^{m}(\gamma_{1i} + \delta_{1i} z_{i:m:n}) - \sum_{i=1}^{m} z_{i:m:n} - \sum_{i=1}^{m} R_i(\alpha_{1i} + \beta_{1i} z_{i:m:n})\right] = 0. \tag{3.6}$$

Upon solving the equations (3.5) and (3.6) for σ and θ, we derive AMLEs of σ and θ as follows;

$$\widetilde{\sigma}_1 = \frac{-B_1 + \sqrt{B_1^2 - 4A_1C_1}}{2A_1} \tag{3.7}$$

and

$$\widetilde{\theta}_1 = M_1\widetilde{\sigma}_1 + M_2, \tag{3.8}$$

where

$$A_1 = 2m - mM_1^2 - \sum_{i=1}^{m} R_i\alpha_{2i} + M_1\sum_{i=1}^{m} R_i\beta_{2i},$$

$$B_1 = 2M_1\sum_{i=1}^{m}(X_{i:m:n} - M_2) - \sum_{i=1}^{m} R_i\beta_{2i}(X_{i:m:n} - M_2),$$

$$C_1 = -\sum_{i=1}^{m}(X_{i:m:n} - M_2)^2, \quad M_1 = A_0/C_0, \quad M_2 = B_0/C_0, \quad A_0 = \sum_{i=1}^{m}(\gamma_{1i} - R_i\alpha_{1i}),$$

$$B_0 = \sum_{i=1}^{m}(\delta_{1i} + 1 - R_i\beta_{1i})X_{i:m:n}, \quad C_0 = \sum_{i=1}^{m}(\delta_{1i} + 1 - R_i\beta_{1i}).$$

Second, making use of the approximate expressions in (3.2) and (3.4), we may approximate the likelihood equation (2.6) as follows;

$$\frac{\partial \ln L}{\partial \sigma} \approx -\frac{1}{\sigma}\left[2m - \sum_{i=1}^{m} z_{i:m:n}^2 - \sum_{i=1}^{m} R_i(\alpha_{1i} + \beta_{1i} z_{i:m:n}) z_{i:m:n}\right] = 0. \tag{3.9}$$

Upon solving the equations (3.9) and (3.6) for σ and θ, we derive another AMLEs of σ and θ as follows;

$$\widetilde{\sigma}_2 = \frac{-B_2 + \sqrt{B_2^2 - 4A_2C_2}}{2A_2} \tag{3.10}$$

and

$$\widetilde{\theta}_2 = M_1\widetilde{\sigma}_2 + M_2, \tag{3.11}$$

where

$$A_2 = 2m - mM_1^2 - M_1^2 \sum_{i=1}^{m} R_i\beta_{1i} + M_1 \sum_{i=1}^{m} R_i\alpha_{1i},$$

$$B_2 = 2M_1 \sum_{i=1}^{m}(X_{i:m:n} - M_2) - \sum_{i=1}^{m} R_i\alpha_{1i}(X_{i:m:n} - M_2) + 2M_1 \sum_{i=1}^{m} R_i\beta_{1i}(X_{i:m:n} - M_2),$$

$$C_2 = -\sum_{i=1}^{m}(X_{i:m:n} - M_2)^2 - \sum_{i=1}^{m} R_i\beta_{1i}(X_{i:m:n} - M_2)^2.$$

4. Least square estimation

The LSE is obtained by minimizing

$$Q(\sigma, \theta) = \sum_{i=1}^{m} \left\{ F(z_{i:m:n}) - \frac{i}{(n+1)} \right\}^2. \tag{4.1}$$

In the case of the double Rayleigh distribution, the equation (4.1) becomes

$$Q(\sigma, \theta) = \sum_{i=1}^{m} \left[\frac{1}{2} \left[1 + sgn(z_{i:m:n}) \left\{ 1 - \exp\left(-\frac{z_{i:m:n}^2}{2} \right) \right\} \right] - \frac{i}{(n+1)} \right]^2. \tag{4.2}$$

To minimize the equation (4.2) with respect to σ and θ, we differentiate with respect to these parameters, which leads to following equations;

$$\sum_{i=1}^{m} \left[\left\{ \frac{1}{2} \left(1 + sgn(z_{i:m:n}) \left(1 - \exp\left(-\frac{z_{i:m:n}^2}{2} \right) \right) \right\} - \frac{i}{(n+1)} \right\} f(z_{i:m:n}) z_{i:m:n} \right] = 0, \tag{4.3}$$

and

$$\sum_{i=1}^{m} \left[\left\{ \frac{1}{2} \left\{ 1 + sgn(z_{i:m:n}) \left(1 - \exp\left(-\frac{z_{i:m:n}^2}{2} \right) \right) \right\} - \frac{i}{(n+1)} \right\} f(z_{i:m:n}) \right] = 0. \tag{4.4}$$

Since the equations (4.3) and (4.4) can not be solved explicitly, the solutions of the equations (4.3) and (4.4) are obtain by using numerical method.

5. Simulated results

From the above formula, the mean squared errors of the estimators are simulated by Monte Carlo method (based on 10,000 Monte Carlo runs) for sample size n =20(10)40, different choices of the effective sample size m, and different progressive censoring schemes with the complete sample in each case.

For simplicity in notation, we will denote the scheme $(0, 0, \ldots, 0, n - m)$ by $((n - m) * 0, n - m)$, for example, (10*0) denotes the progressive censoring scheme $(0, 0, \ldots, 0)$ and (3 * $0, 2, 2, 0)$ denotes the progressive censoring scheme $(0, 0, 0, 2, 2, 0)$.

Table 5.1 The relative mean squared errors of the estimators of the scale parameter σ.

n	m	scheme	MSE			
			$\widetilde{\sigma}_1$	$\widetilde{\sigma}_2$	$\widehat{\sigma}$	σ^*
20	20	(20*0)	0.013858	0.013858	0.040966	0.195462
	16	(1*0,2,2*0,2,11*0)	0.019801	0.020157	0.070427	0.213874
	16	(1*0,2,1*0,2,12*0)	0.01797	0.018141	0.067566	0.211235
	12	(8,11*0)	0.022502	0.022911	0.077943	0.201898
	12	(2*0,1,3*0,4,2*0,3,2*0)	0.039769	0.037711	0.190938	0.261847
	10	(5,2*0,5,6*0)	0.03164	0.030842	0.18118	0.235333
	10	(2*0,1,1*0,2,1*0,2,2*0,5)	0.088429	0.08749	0.251845	0.261342
30	30	(30*0)	0.008904	0.008904	0.034516	0.191857
	25	(2,*0,5,22*0)	0.010364	0.010325	0.049456	0.204497
	25	(24*0,5)	0.011738	0.011351	0.059019	0.289563
	20	(3*0,5,3*0,5,16*0)	0.015434	0.015619	0.115466	0.251243
	20	(2*0,10,17*0)	0.012943	0.013067	0.079202	0.217805
	15	(5,6*0,10,7*0)	0.024299	0.0231	0.229261	0.27775
	15	(10,6*0,5,7*0)	0.020343	0.020232	0.155818	0.251158
40	40	(40*0)	0.006598	0.006598	0.03171	0.188425
	35	(12*0,5,22*0)	0.007525	0.007521	0.055631	0.215214
	35	(34*0,5)	0.007828	0.007717	0.04699	0.304348
	30	(6*0,5,8*0,5,14*0)	0.010023	0.009989	0.089985	0.248856
	30	(18*0,10,11*0)	0.010519	0.010897	0.106545	0.271409
	25	(6*0,5,3*0,5,9*0,5,4*0)	0.016746	0.016284	0.147769	0.303203
	25	(1*0,5,12*0,10,10*0)	0.012947	0.013102	0.145874	0.281542
	20	(3*0,3,3*0,10,5*0,3,4*0,4,1*0)	0.026781	0.021881	0.248587	0.313111

Table 5.2 The relative mean squared errors of the estimators of the location parameter θ.

n	m	scheme	MSE			
			$\widetilde{\theta}_1$	$\widetilde{\theta}_2$	$\widehat{\theta}$	θ^*
20	20	(20*0)	0.18304	0.18304	0.142105	0.170457
	16	(1*0,2,2*0,2,11*0)	0.355307	0.355861	0.138405	0.192527
	16	(1*0,2,1*0,2,12*0)	0.273865	0.274382	0.13774	0.191807
	12	(8,11*0)	0.263966	0.265259	0.138804	0.208349
	12	(2*0,1,3*0,4,2*0,3,2*0)	0.478535	0.484173	0.163849	0.34859
	10	(5,2*0,5,6*0)	0.275835	0.275475	0.150914	0.257285
	10	(2*0,1,1*0,2,1*0,2,2*0,5)	0.418522	0.418311	0.235474	0.427277
30	30	(30*0)	0.13926	0.13926	0.141659	0.140011
	25	(2,*0,5,22*0)	0.173404	0.173416	0.139081	0.15606
	25	(24*0,5)	0.160072	0.160387	0.147059	0.224051
	20	(3*0,5,3*0,5,16*0)	0.219111	0.219391	0.139408	0.192052
	20	(2*0,10,17*0)	0.325958	0.326433	0.141579	0.171829
	15	(5,6*0,10,7*0)	0.328323	0.328657	0.148497	0.287855
	15	(10,6*0,5,7*0)	0.20088	0.200797	0.144955	0.23768
40	40	(40*0)	0.117083	0.117083	0.139018	0.116282
	35	(12*0,5,22*0)	0.164357	0.164388	0.13689	0.130908
	35	(34*0,5)	0.124466	0.124557	0.140377	0.151715
	30	(6*0,5,8*0,5,14*0)	0.161412	0.161368	0.136683	0.156025
	30	(18*0,10,11*0)	0.177064	0.176967	0.138431	0.202713
	25	(6*0,5,3*0,5,9*0,5,4*0)	0.296627	0.297357	0.155225	0.274176
	25	(1*0,5,12*0,10,10*0)	0.20664	0.206594	0.139796	0.232196
	20	(3*0,3,3*0,10,5*0,3,4*0,4,1*0)	0.328524	0.333154	0.17438	0.381719

From Table 5.1 and Table 5.2, we have the following results;

As expected, the MSE of all estimators decrease as sample size n increases. For fixed sample size, the MSE increases generally as the number of unobserved or missing data $n-m$ increases.

The performances of the AMLEs $\widetilde{\sigma}_1$ and $\widetilde{\sigma}_2$ are very similar in all aspects. But the AMLEs $\widetilde{\sigma}_1$ and $\widetilde{\sigma}_2$ are more efficient than the MLE $\widehat{\sigma}$ and the LSE $\sigma*$. The estimator $\widetilde{\sigma}_2$ is generally more efficient than $\widetilde{\sigma}_1$ for large $n-m$. Since the MLE $\widehat{\sigma}$ and the LSE $\sigma*$ cannot be solved explicitly, some numerical method must be employed. However, the AMLEs $\widetilde{\sigma}_1$ and $\widetilde{\sigma}_2$ are explicit estimators and have good performance.

For the location parameter, the MLE $\widehat{\theta}$ is generallymore efficient than the others in the sense of the MSE. But the AMLEs $\widetilde{\theta}_1$ and $\widetilde{\theta}_2$ aremore efficient than the others when censoring schemes are (30*0) and (40*0).

References

Balakrishnan, N. (1989). Approximate MLE of the scale parameter of the Rayleigh distribution with censoring. *IEEE Transactions on Reliability*, **38**, 355-357.

Dattatreya Rao, A. V. and Narasimham, V. L. (1989). Linear estimation in double Weibull distribution. *Sankhya: The Indian Journal of Statistic*, **51**, 24-64.

Han, J. T. and Kang, S. B. (2006). Estimation for two-parameter Rayleigh distribution based on multiply Type-II censored sample. *Journal of the Korean Data & Information Science Society*, **17**, 1319-1328.

Han, J. T. and Kang, S. B. (2008). Estimation for the double Rayleigh distribution based on multiply Type-II censored samples. *Communications of the Korean Statistical Society*, **15**, 367-378.

Kang, S. B. (2007). Estimation in a half-triangle distribution based on multiply Type-II censored samples. *Journal of the Korean Data & Information Science Society*, **18**, 793-801.

Kang, S. B., Lee, H. J. and Han, J. T. (2004). Estimation of Weibull scale parameter based on mutiply Type-II censored samples. *Journal of the Korean Data & Information Science Society*, **15**, 593-603.

Lee, H. J., Han, J. T. and Kang, S. B. (2008). Estimation for a triangular distribution based on multiply Type-II censored samples. *Journal of the Korean Data & Information Science Society*, **19**, 319-330.

Son, H. and Woo, J. (2007). Estimating a skewed parameter and reliability in a skew-symmetric double Rayleigh distribution. *Journal of the Korean Data & Information Science Society*, **18**, 1205-1214.

Vasudeva R, A., Dattatreya Rao, A. V. and Narasimham, A. L. (1991). Optimum linear unbiased estimation of the scale parameter by absolute values of order statistics in the double exponential and the double Weibull distributions. *Communications in Statistics-Simulation and Computation*, **20**, 1139-1158.