Journal of Modern Applied Statistical Methods

Volume 13 | Issue 1 Article 17

5-1-2014

Estimation and Testing in Type-II Generalized Half Logistic Distribution

R R. L. Kantam
Acharya Nagarjuna University, India, kantam.rrl@gmail.com

V Ramakrishna
K. L. University, Vaddeswaram, India, vramakrishna2006@gmail.com

M S. Ravikumar
Acharya Nagarjuna University, India, msrk.raama@gmail.com

Journal of Modern Applied Statistical Methods
May 2014, Vol. 13, No. 1, 267-277.

ISSN 1538 − 9472

Estimation and Testing in Type-II Generalized Half Logistic Distribution

R. R. L. Kantam
Acharya Nagarjuna University
Nagarjunanagar, India

V. Ramakrishna
K. L. University
Vaddeswaram, India

M. S. Ravikumar
Acharya Nagarjuna University
Nagarjunanagar, India

A generalization of the Half Logistic Distribution is developed through exponentiation of its survival function and named the Type II Generalized Half Logistic Distribution (GHLD). The distributional characteristics are presented and estimation of its parameters using maximum likelihood and modified maximum likelihood methods is studied with comparisons. Discrimination between Type II GHLD and exponential distribution in pairs is conducted via likelihood ratio criterion.

Keywords: Generalized Half Logistic Distribution (GHLD), maximum likelihood estimation (MLE), modified maximum likelihood estimation (MMLE), mean square error (MSE), likelihood ratio type criterion, percentiles, power of the test

Introduction

In life testing and reliability studies a combination of monotone and constant failure rates over various segments of the range of lifetime of a random variable is also known as bath tub or non-monotone failure rate. In biological and engineering sciences, situations of non-monotone failure rates are common (see Rajarshi & Rajarshi (1988) for a comprehensive narration of these models). Mudholkar, et al. (1995) presented an extension of the Weibull family that contains unimodel distributions with bathtub failure rates and also allows for a broader class of monotone hazard rates. They named their extended version the Exponentiated Weibull Family.

Gupta and Kundu (1999) proposed a new model called the generalized exponential distribution. If θ is a positive real number and $F(x)$ is the cumulative

Dr. R. R. L. Kantam is a Professor in the Department of Statistics. Email him at kantam.rrl@gmail.com. V. Ramakrishna is an Associate Professor in the Department of Computer Science and Engineering. Email him at: vramakrishna2006@gmail.com. M. S. Ravikumar is a UGC Research Fellow in the Department of Statistics. Email him at: msrk.raama@gmail.com.

distribution function (cdf) of a continuous positive random variable, then $[F(x)]^{\theta}$ and the corresponding probability distribution may be termed an exponentiated or generalized version of $F(x)$.

A half logistic model obtained as the distribution of absolute standard logistic variate is a probability model of recent origin (Balakrishnan, 1985). Its standard probability density function, cumulative distribution function and hazard functions are given by:

$$f(x) = \frac{2e^{-x}}{(1+e^{-x})^2},\ x \geq 0 \tag{1}$$

$$F(x) = \left[\frac{1-e^{-x}}{1+e^{-x}}\right],\ x \geq 0 \tag{2}$$

$$F(x) = \left[\frac{1-e^{-x}}{1+e^{-x}}\right], x \geq 0\,. \tag{3}$$

Kantam et al. (2011) adopted this generalization to the well-known half logistic distribution, and named it the Type-I Generalized Half Logistic Distribution (GHLD).

Consider a series system of θ components with individually and identically distributed (iid) individual lifetimes, for example, $F(x)$. The reliability function of such a system is given by $[1 - F(x)]^{\theta}$; hence, the distribution function of the lifetime random variable of a series system is $1 - [1 - F(x)]^{\theta}$.

Taking $F(x)$ as the half logistic model given by Equation (2), the corresponding distribution is termed the Type-II Generalized Half Logistic Distribution (GHLD-II). Its pdf, cdf and hazard function are given by:

$$f(x) = \frac{\theta(2e^{-x})^{\theta}}{(1+e^{-x})^{\theta+1}}, x > 0,\ \theta > 0 \tag{4}$$

$$F(x) = 1 - \left[\frac{2e^{-x}}{1+e^{-x}}\right]^{\theta}, x > 0,\ \theta > 0 \tag{5}$$

$$h(x) = \frac{\theta}{1+e^{-x}}, x > 0,\ \theta > 0\,. \tag{6}$$

Balakrishnan and Sandhu (1995) suggested a new probability model with a standard pdf and cdf given by:

$$f(x) = \frac{2(1-kx)^{(1/k)-1}}{[1+(1-kx)^{1/k}]^2}, 0 \le x \le \frac{1}{k}, k \ge 0 \tag{7}$$

$$F(x) = \frac{1-(1-kx)^{1/k}}{1+(1-kx)^{1/k}}, 0 \le x \le \frac{1}{k}, k \ge 0 . \tag{8}$$

The limits of (7) and (8) as $k \to \infty$ are respectively (1) and (2) – the pdf and cdf of HLD. Balakrishnan and Sandhu (1995) called the distribution (7) and (8) Generalized HLD.

Olapade (2008) considered two distributions and discussed their distributional properties, order statistics in samples from these distributions: He named these distributions type-I and type-III GHLD, respectively. The types of generalized HLD of Olapade (2008) are through truncation of the type-I and type-III generalized logistic distributions from Balakrishnan and Leung (1988) at the origin. Thus, this type-II GHLD is conceptually different from the GHLDs of Balakrishnan and Sandhu (1995) and Olapade (2008). Hence, the proposed models motivated a separate research study.

Estimation in Type-II Generalized Half Logistic Distribution (GHLD-II)

The probability density function and distribution function of GHLD-II with scale parameter σ and power parameter θ are given by:

$$f(x) = \frac{\theta(2e^{-x/\sigma})^{\theta}}{\sigma(1+e^{-x/\sigma})^{\theta+1}}, \; 0 < x < \infty, \; \sigma > 0, \; q > 0 \tag{9}$$

$$F(x) = 1 - \left[\frac{2e^{-x/\sigma}}{1+e^{-x/\sigma}}\right]^{\theta}, \; 0 < x < \infty, \; \sigma > 0, \; q > 0 . \tag{10}$$

Let $x_1 < x_2 < \ldots < x_n$ be an ordered sample of size n from GHLD-II. The log likelihood function of the sample is

$$\log L = n(\log\theta - \log\sigma) + \sum_{i=1}^{n}\left[\theta\log 2 - \theta\frac{x_i}{\sigma} - \theta\log(1+e^{-x_i/\sigma}) - \log(1+e^{-x_i/\sigma})\right]$$

The log likelihood equations to estimate the parameters σ and θ are given by

$$\frac{\partial\log L}{\partial\sigma} = 0, \ \frac{\partial\log L}{\partial\theta} = 0,$$

$$\frac{\partial\log L}{\partial\sigma} = 0 \Rightarrow \sum_{i=1}^{n}\frac{x_i}{\sigma}\left[\frac{\theta - e^{-x_i/\sigma}}{1+e^{-x_i/\sigma}}\right] = n \tag{11}$$

$$\frac{\partial\log L}{\partial\theta} = 0 \Rightarrow \theta = \frac{n}{\sum_{i=1}^{n}\log(1+e^{-x_i/\sigma}) + \frac{1}{\sigma}\sum_{i=1}^{n}x_i - \log 2^n} \tag{12}$$

It can be seen that these two equations must be solved iteratively for θ and σ for a given sample. The asymptotic variances and covariances of MLEs of σ and θ can be obtained by inverting the information matrix whose elements are the mathematical expectation of the following expressions:

$$-\left(\frac{\partial^2\log L}{\partial\sigma^2}\right) = \sum_{i=1}^{n}\left[\frac{x_i^2}{\sigma^4}\frac{(1-\theta+2e^{-x_i/\sigma})e^{-x_i/\sigma}}{(1+e^{-x_i/\sigma})^2}\right] + \frac{n}{\sigma^2} \tag{13}$$

$$-\left(\frac{\partial^2\log L}{\partial\theta^2}\right) = \frac{n}{\theta^2} \tag{14}$$

$$-\left(\frac{\partial^2\log L}{\partial\theta\partial\sigma}\right) = -\frac{1}{\sigma^2}\sum_{i=1}^{n}\left[\frac{x_i}{(1+e^{-x_i/\sigma})}\right] \tag{15}$$

These equations, evaluated at estimates of θ and σ, provide am estimated dispersion matrix. In order to obtain an analytical estimator for σ, its estimating equation is approximated by some admissible expression.

Equation (11) to get MLE of σ, after simplification would become

$$\sum_{i=1}^{n} \frac{z_i(\theta - e^{-z_i})}{1+e^{-z_i}} - n = 0 \text{ where } z_i = \frac{x_i}{\sigma} \tag{16}$$

To obtain the analytical expression for σ, approximate the following expression in (16) by some linear function in the corresponding population quartile. Let,

$$G(z_i) = \frac{z_i(\theta - e^{-z_i})}{(1+e^{-z_i})} \tag{17}$$

approximate

$$G(z_i) \approx \alpha_i + \beta_i z_i \tag{18}$$

where α_i , β_i are to be suitably found. After using this approximation in (16) the solution for σ is

$$\hat{\sigma} = \frac{\sum_{i=1}^{n} \beta_i x_i}{n - \sum_{i=1}^{n} \alpha_i} \tag{19}$$

This estimator is named the MMLE of σ, which is a linear estimator in x_i's

To obtain α_i, β_i, let $p_i = \dfrac{i}{n+1}; i = 1, 2, ..., n$ and let t_i, t_i^* be the solutions of equations:

$$F(t_i) = p_j - \sqrt{\frac{p_i q_i}{n}} = p_i^{'} \text{ (for example)} \tag{20}$$

$$F(t_i^*) = p_i + \sqrt{\frac{p_i q_i}{n}} = p_i^{''} \text{ (for example), where } q = 1 - p_i \tag{21}$$

where $F(.)$ is cdf of GHLD-II.

The intercept α_i and slope β_i of linear approximation in the Equation (18) are respectively given by

$$\beta_i = \frac{G(t_i^*) - G(t_i)}{t_i^* - t_i} \tag{22}$$

$$\alpha_i = G(\mathrm{t_i^*}) - \beta_i t_i^*. \tag{23}$$

Using distribution function $F(.)$ of GHLD-II, the expressions for t_i, t_i^* are given by

$$t_i = \log\left(\frac{2-(1-p_t^{'})^{1/\theta}}{(1-p_t^{'})^{1/\theta}}\right),\ t_i^* = \log\left(\frac{2-(1-p_t^{''})^{1/\theta}}{(1-p_t^{''})^{1/\theta}}\right).$$

Table 1 shows the values of α_i, β_i for various θ and n. The MMLE of σ can be shown to be equivalent to the exact MLE with respect to the asymptotic variance. Their performance in small samples is also studied through simulation because the exact MLE is an iterative solution. The empirical sample characteristics are given in Table 2, which indicates the following:

1. The empirical sample characteristics bias, variance and MSE decrease as sample size increases.

2. MMLE is generally more biased than MLE; with reference to variance as well as MSE, MMLE is better than MLE for small samples.

Table 1. Intercept and Slope of the Approximation $G(Z_i) = \alpha_i + \beta_i z_i$ (GHLD –II)

		$\theta = 2$		$\theta = 3$		$\theta = 4$	
n	i	α_i	β_i	α_i	β_i	α_i	β_i
5	1	0.0000	0.7752	0.0000	1.2528	0.0000	1.7410
	2	-0.0612	1.0780	-0.0391	1.5442	-0.0286	2.0251
	3	-0.2071	1.4122	-0.1407	1.8906	-0.1060	2.3751
	4	-0.4780	1.7832	-0.3573	2.3293	-0.2823	2.8464
	5	-0.4013	1.5789	-0.3942	2.2353	-0.3891	2.8947
10	1	0.0000	0.6432	0.0000	1.1294	0.0000	1.6223
	2	-0.0150	0.7934	-0.0092	1.2679	-0.0066	1.7545
	3	-0.0477	0.9512	-0.0299	1.4170	-0.0216	1.8985
	4	-0.1010	1.1170	-0.0650	1.5785	-0.0476	2.0569
	5	-0.1789	1.2912	-0.1187	1.7552	-0.0883	2.2334
	6	-0.2866	1.4743	-0.1974	1.9509	-0.1496	2.4334
	7	-0.4310	1.6668	-0.3111	2.1713	-0.2413	2.6662
	8	-0.6219	1.8683	-0.4780	2.4268	-0.3826	2.9484
	9	-0.8720	2.0755	-0.7370	2.7393	-0.6201	3.3206
	10	-0.5324	1.7681	-0.5479	2.4572	-0.5547	3.1336
15	1	0.0000	0.5960	0.0000	1.0870	0.0000	1.5020
	2	-0.0066	0.6969	-0.0040	1.1781	-0.0028	1.6684
	3	-0.0207	0.8003	-0.0127	1.2736	-0.0091	1.7596
	4	-0.0430	0.9072	-0.0268	1.3739	-0.0193	1.8563
	5	-0.0744	1.0176	-0.0472	1.4797	-0.0343	1.9591
	6	-0.1161	1.1318	-0.0749	1.5917	0.0550	2.0691
	7	-0.1698	1.2498	-0.1116	1.7106	0.0827	2.1875
	8	-0.2357	1.3719	-0.1589	1.8377	-0.1191	2.3158
	9	-0.3171	1.4981	-0.2196	1.9744	-0.1668	2.4564
	10	-0.4159	1.6286	-0.2972	2.1226	-0.2292	2.6121
	11	-0.5350	1.7632	-0.3970	2.2852	-0.3120	2.7877
	12	-0.6781	1.9015	-0.5274	2.4665	-0.4245	2.9907
	13	-0.8492	2.0417	-0.7032	2.6737	-0.5843	3.2348
	14	-1.0496	2.1780	-0.9559	2.9223	-0.8344	3.5533
	15	-0.5925	1.8548	-0.6260	2.5684	-0.6430	3.2610

Table 1, continued

		$\theta = 2$		$\theta = 3$		$\theta = 4$	
n	i	α_i	β_i	α_i	β_i	α_i	β_i
20	1	0.0000	0.5732	0.0000	1.0656	0.0000	1.5617
	2	0.0037	0.6482	-0.0022	1.1334	0.0016	1.6259
	3	0.0115	0.7251	-0.0070	1.2037	0.0050	1.6927
	4	-0.0237	0.8039	-0.0146	1.2766	-0.0104	1.7623
	5	-0.0406	0.8847	-0.0252	1.3522	-0.0182	1.8351
	6	-0.0627	0.9675	-0.0394	1.4309	-0.0285	1.9112
	7	-0.0903	1.0525	-0.0575	1.5129	-0.0419	1.9913
	8	-0.1241	1.1395	-0.0802	1.5986	0.0588	2.0755
	9	-0.1645	1.2288	-0.1080	1.6883	-0.0798	2.1647
	10	-0.2123	1.3204	-0.1417	1.7825	-0.1057	2.2593
	11	-0.2683	1.4143	-0.1825	1.8818	-0.1373	2.3603
	12	-0.3332	1.5107	-0.2314	1.9869	-0.1760	2.4686
	13	-0.4084	1.6094	-0.2904	2.0987	-0.2234	2.5859
	14	-0.4948	1.7105	-0.3615	2.2183	-0.2819	2.7138
	15	-0.5942	1.8139	-0.4479	2.3474	-0.3548	2.8551
	16	-0.7081	1.9192	-0.5543	2.4881	-0.4475	3.1039
	17	-0.8383	2.0254	-0.6878	2.6439	-0.5687	3.1967
	18	-0.9857	2.1303	-0.8609	2.8203	-0.7349	3.4153
	19	-1.1453	2.2265	-1.0994	3.0286	-0.9857	3.6982
	20	-0.6282	1.9063	-0.6757	2.6416	-0.7012	3.3450

Table 2. Empirical Sample Characteristics (Type-II GHLD)

		Bias		Variance		MSE	
θ	n	MLE	MMLE	MLE	MMLE	MLE	MMLE
2	5	0.1077	0.0910	0.0651	0.0121	0.0766	0.0203
	10	0.0551	0.0659	0.0320	0.0079	0.0350	0.0122
	15	0.0364	0.0522	0.0206	0.0060	0.0219	0.0087
	20	0.0273	0.0427	0.0153	0.0048	0.0160	0.0066
3	5	0.1064	0.0977	0.0643	0.0125	0.0756	0.0220
	10	0.0549	0.0667	0.0320	0.0081	0.0350	0.0125
	15	0.0364	0.0530	0.0207	0.0061	0.0220	0.0089
	20	0.0274	0.0435	0.0154	0.0049	0.0161	0.0067
4	5	0.1055	0.0926	0.0636	0.0131	0.0747	0.0216
	10	0.0546	0.0676	0.0318	0.0085	0.0347	0.0130
	15	0.0362	0.0538	0.0206	0.0064	0.0219	0.0092
	20	0.0273	0.0443	0.0153	0.0051	0.0160	0.0070

GHLD-II vs. Exponential Model

The discrimination between GHLD-II and the exponential model is made using the likelihood ratio (LR) criterion. Specify GHLD-II as null population (P_0) and the exponential model as alternative population (P_1). A null hypothesis is proposed as H_0: a given sample belongs to GHLD-II (P_0) versus an alternative hypothesis H_1: the sample belongs to the population Exponential model (P_1). Let L_1, L_0, respectively, stand for the likelihood function of a sample with population P_1 and P_0. The percentiles of the LR criterion L1/L0 are obtained by simulation as:

10,000 random samples of sizes n = 5, 10, 15, 20 are generated from the null population P_0 and its parameters are estimated using each sample. The value of the likelihood function of the null population is computed at the generated sample observations and the corresponding parameter estimates; this value is denoted by L_0. Using the same sample, generated from P_0, the parameters and likelihood function value of the alternative population are calculated, for example, L_1. The values of L_1/L_0 over 10,000 runs are sorted and selected percentiles are identified for a given n, θ (see Table 3).

Table 3. Percentiles of L_1/L_0 (P_0 : GHLD-II, P_1: Exponential)

θ	$n \setminus p$	0.00135	0.01	0.025	500	0.95	0.975	0.99865
2	5	0.7468	0.9743	1.3335	1.7250	2.5433	2.6067	4.4061
	10	0.4786	0.7651	1.2327	1.6918	4.7663	4.8496	6.6528
	15	0.3369	0.7567	1.1770	1.6473	6.0976	7.4550	8.6546
	20	0.2520	0.7344	1.0456	1.5327	8.9127	8.9845	10.7528
3	5	1.6877	1.9325	2.2646	2.6432	3.4623	4.6379	20.6042
	10	2.5396	3.0615	4.0111	5.0243	8.7750	9.0357	39.9667
	15	3.1753	5.4376	7.9391	9.8175	18.5364	18.7628	50.6341
	20	3.9089	9.7390	14.7436	19.4296	54.4206	69.0316	80.0497
4	5	3.6630	4.5150	5.1328	4.0879	18.2493	18.4501	73.5894
	10	10.1778	12.7498	16.3453	12.4361	30.6046	31.1481	81.5585

The entries under the column headings 0.95 in Table 3 may be taken as 5% level of significance critical values for discriminating between the GHLD-II and exponential models. The powers of the test statistic L_1/L_0 are also evaluated through simulation by calculating L_1/L_0 with samples generated from exponential population (P_1) and estimating, the parameters calculating the values of the likelihood functions L_1, L_0 with sample from P_1. The proportion of L_1/L_0 values falling above 95th percentile of L_1/L_0 would become the power of the LR test criterion (see Table 4). It is observed that the discrimination between GHLD-II and exponential models falls with increased sample size, indicating less distinguishability between the exponential model and GHLD-II.

Table 4. Powers of LR Test Criterion at α = 0.05

θ	n \ Distributions	GHLD-II vs. Exponential
2	5	0.9123
	10	0.9239
	15	0.9373
	20	0.9441
3	5	0.9135
	10	0.9159
	15	0.9176
	20	0.9161
4	5	0.9072
	10	0.9057
	15	0.9053
	20	0.9025

References

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

Balakrishnan, N., & Leung, M. Y. (1988). Order statistics from type-I generalized logistic distribution. *Communication Statistics – Simulation and Computing*, 17(1): 25-50.

Balakrishnan, N., & Sandhu, R. A. (1995). Recurrence relations for single and product moments of order Statistics from a generalized half logistic distribution with applications to inference. *Journal of Statistical Computation and Simulation*, *52*(4): 385-398.

Gupta, R. D., & Kundu, D. (1999). Generalized exponential distributions. *Australian and New Zealand Journal of Statistics*, *41*: 173-188.

Kantam, R. R. L., & Srinivasa Rao, G. (1993). Reliability estimation in Rayleigh distribution with censoring some approximations to ML Method. *Proceedings of II Annual Conference of Society for Development of Statistics, Acharya Nagarjuna University*: 56-63.

Kantam, R. R. L., & Srinivasa Rao, G. (2002). Log-logistic distribution: Modified Maximum likelihood estimation. *Gujarat Statistical Review*, *29*(1-2): 25-36.

Kantam, R. R. L., and Sriram, B. (2003). Maximum likelihood estimation from censored samples: Some modifications in length biased version of exponential model. *Statistical methods*, *5*(1): 63-78.

Mehrotra, K. G., & Nanda, P. (1974). Unbiased estimation of parameters by order statistics in the case of censored samples. *Biometrika, 61*: 601-606.

Mudholkar, G. S., Srivastava, D., & Freimer, M. (1995). Exponentiated Weibull family: A reanalysis of the bus-motor failure data. *Technometrics*, *37*(4): 436-445.

Olapade, A. K. (2008). On Type III Generalized Half Logistic Distribution. arXiv:0806.1580v1 [math.ST] 10 Jun 2008.

Pearson, T., & Rootzen, H. (1977). Simple highly efficient estimators for a type-I Censored Normal sample, *Biometrika*, *64*: 123-128.

Rajarshi, S., & Rajarshi, M. B. (1988). Bathtub distributions: A review. *Communication in Statistics –Theory & Methods*, *17*: 2597-2621.

Rosaiah, K., Kantam, R. R. L., & Narasimham, V. L. (1993a). ML and Modified ML Estimation in gamma distribution with known prior relation among the Parameters. *Pakistan Journal of Statistics*, *9*(3)B: 37-48.

Rosaiah, K., Kantam, R. R L., & Narasimham, V. L. (1993b). On modified maximum likelihood estimation of gamma parameters. *Journal of Statistical Research*, *27*(1-2): 15-28.

Tiku, M. L. (1967). Estimating the mean and standard deviation from a censored Normal sample. *Biometrika*, *54*: 155-165.

Tiku, M. L., & Suresh, R. P. (1992). A new method of estimation for location and scale parameters. *Journal of Statistical Planning and Inference*, 30: 281-92.