The Korean Communications in Statistics
Vol. 13 No. 3, 2006, pp.503-512

Kernel Inference on the Inverse Weibull Distribution

M. Maswadah[1)]

Abstract

In this paper, the Inverse Weibull distribution parameters have been estimated using a new estimation technique based on the non-parametric kernel density function that introduced as an alternative and reliable technique for estimation in life testing models. This technique will require bootstrapping from a set of sample observations for constructing the density functions of pivotal quantities and thus the confidence intervals for the distribution parameters. The performances of this technique have been studied comparing to the conditional inference on the basis of the mean lengths and the covering percentage of the confidence intervals, via Monte Carlo simulations. The simulation results indicated the robustness of the proposed method that yield reasonably accurate inferences even with fewer bootstrap replications and it is easy to be used than the conditional approach. Finally, a numerical example is given to illustrate the densities and the inferential methods developed in this paper.

Keywords : Kernel density estimation; conditional inference; covering percentage.

1. Introduction

The purpose of this paper is to introduce a new parameter estimation procedure in statistical inference for making inference on lifetime parameters or function thereof directly from the data without any prior assumptions about the underlying failure model parameters. This technique uses the non-parametric kernel density estimation that asymptotically converges to any density function depending only on a random sample, though the underlying distribution is not known and requires fewer bootstrap replications to attain any level of accuracy. These properties make the kernel estimation approach is a quite general and applicable to any problem and that was the merit for using this function as a tool for estimation.

The statistical performances of the proposed kernel procedure have been compared, via Mote Carlo simulation, to the performances of the classical conditional inference based on the covering percentage and the mean lengths of

1) Department of Mathematics, Faculty of Science, South Valley University, Aswan, Egypt.
E-mail : maswadah@hotmail.com

the confidence intervals for the parameters. The simulation results indicated quite improvement and accurate results, even when the number of bootstrap replications is extremely small comparing to the conditional inference.

In this study, the kernel estimation has been applied for deriving the unconditional distributions for pivotal quantities of the parameters with comparing to those conditional distributions based on the classical conditional inference when the experimental data are collected under complete samples from the Inverse Weibull distribution (IWD) that has probability density function (pdf) given by :

$$f(x) = \alpha\beta^{-\alpha}x^{-(\alpha+1)}\exp[-(x\beta)^{-\alpha}],\ x > 0, \tag{1.1}$$

where $\alpha(>0)$ and $\beta(>0)$ are the shape and scale parameters respectively.

This distribution has been used extensively as a model in the analysis of life testing data. The statistical inference for the IWD has been investigated by several authors such as Calariba and Pulcini (1990, 1992) have been derived the maximum likelihood estimators and the Bayes estimators for the parameters based on complete and censored samples. Maswadah (2003) derived the conditional confidence intervals for the parameters based on the generalized order statistics. For a detailed discussion on various properties and uses of this distribution, see Johnson et al. (1995).

2. Kernel Function

2.1 Basic Definitions

In this section, the basic elements associated with the kernel estimators of the density function are presented, which has been extensively studied see, for example Guillamon et al. (1998, 1999). Also a good discussion for the kernel estimation techniques can be found in Scott (1992). In the univariate case, the general kernel estimator based on a random sample $x_1, x_2, x_3, \cdots, x_n$ of size n from the random variable X with unknown probability density function $f(x)$and support on $(0,\infty)$ is given by

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n}K\left(\frac{x-x_i}{h}\right) \tag{2.1}$$

where h is called the bandwidth or smoothing parameter which chosen such that $h\to 0$ and $nh\to\infty$ as $n\to\infty$. The role of the bandwidth is to scale for our kernels, if it is large the density estimate could be too smooth, otherwise the estimate could be too variable. Unfortunately, the choice of h is the main problem of the kernel, where the optimal one is not known in general, it has been investigated by several authors such as Terrell (1990) and Jones (1991) they concluded that, for a large amount of data, the mean integrated squared error of $\hat{f}(x)$ is minimized when $h = C\hat{\sigma}n^{-0.2}$, where the estimated value of the population standard deviation $\hat{\sigma}$ could be used with different choices: Firstly, it could be used as the sample standard deviation, and $C = 1.06$ see Abramson (1982). Secondly, it could be used as the inter-quintile range which is defined as $\hat{R} = X_{[.75n]} - X_{[.25n]}$, and $C = 0.79$. Finally, it could be used as $A = \min(S, \hat{R}/1.34)$, and $C = 0.9$. It is worthwhile to say that these different values of h give approximate results , however the optimal one is $h = 1.059\hat{\sigma}\, n^{-0.2}$, which is used in our simulations where $\hat{\sigma}$ is the sample standard deviation.

The kernel function K is a symmetric probability density function on the entire line and satisfies the following conditions:

$$\int K(u)du = 1 \ , \int uK(u)du = 0 \ \text{and} \ \int u^2K(u)du < \infty .$$

The role of K is to spread out the contribution of each data point in our estimate of the parent distribution and the estimate $\hat{f}(x)$ is bin-independent regardless of the choice of K. Though there are variety of kernel functions with different properties have been used in the literature, but an obvious and natural choice of K is the standard Gaussian kernel, for its continuity, differentiability, and locality properties.

2.2 Kernel Estimates

We propose a simple and tractable algorithm for estimating the density functions of the pivotal quantities based on the kernel estimate as the following:

1. Let $x_1, x_2, x_3, \cdots, x_n$ be a random sample of size n from the random variable$X$, whose pdf is $f(x;\theta)$, where θ represents the unknown parameter with support on $(0, \infty)$.
2. Bootstrapping with replacement n samples $X_1^*, X_2^*, X_3^*, \cdots, X_n^*$ of size n from the given random sample in step 1, where $X_i^* = (x_1^*, x_2^*, x_3^*, \cdots, x_n^*)$ for $i = 1, 2, \cdots, n$.
3. For each sample in step 2, calculate the pivotal quantity Z for the parameter θ based on its maximum likelihood estimator. Thus we have an objective and informative random sample $Z_1, Z_2, Z_3, \cdots, Z_n$ of size n, which constitute the sampling distribution of the pivotal Z .
4. Finally, based on the informative sample in step 3 we can use the kernel estimator (2.1) for estimating $g(z)$ at any given value for the pivotal Z thus the probability interval estimates for the pivotal and the confidence interval for the unknown parameter θ can be derived.

Utilizing the above algorithm, the kernel estimator of the quantile Z_p, of order p, for Z can be derived as:

$$G(z_p)=\int_0^z \hat{g}(z)dz=\frac{1}{nh}\sum_{i=1}^{n}\int_0^z K\left(\frac{z-z_i}{h}\right)dz=p$$

Thus

$$\sum_{i=1}^{n}\prod\left(-\frac{z_i}{h},\frac{z_p-z_i}{h}\right)=np, \tag{2.2}$$

where

$$\prod(x_1, x_2)=\int_{x_1}^{x_2}K(y)dy.$$

For deriving the value of the quantile estimator Z_p, equation (2.2) can be solved recurrently as the limit of the sequence $\{\widetilde{Z}_1, \widetilde{Z}_2, \widetilde{Z}_3, \cdots\}$ that defined by the formulas

$$\widetilde{Z}_1=\frac{1}{n}\sum_{i=1}^{n}Z_i ,\widetilde{Z}_{r+1}=\widetilde{Z}_r+C\left[np-\sum_{i=1}^{n}\prod\left(-\frac{Z_i}{h},\frac{\widetilde{Z}_r-Z_i}{h}\right)\right] \text{ for } r=1,2,3,\cdots \tag{2.3}$$

The convergence of (2.3) is guaranteed by the condition $0< C\le \dfrac{2h}{nL_1}$ where $L_1=K(0)$, see Kulczycki (1999).

The central rule for applying this technique is deriving the MLEs of the IWD parameters $\theta=(\alpha, \beta)$ based on the complete sample, which are the solutions of the two equations:

$$\beta=\left(\sum_{i=1}^{n}x_i^{-\alpha}/n\right)^{1/\alpha} \tag{2.4}$$

$$\frac{n}{\alpha}-\sum_{i=1}^{n}\ln(x_i)+n\sum_{i=1}^{n}x_i^{-\alpha}\ln(x_i)/\sum_{i=1}^{n}x_i^{-\alpha}=0 \tag{2.5}$$

Using an iterative technique such as Newton-Raphson method for solving (2.5), we can derive the MLE for α and then for β from (2.4).

3. Conditional Inference

In this section we outline the key ideas for deriving the conditional confidence intervals for parameters of the IWD distribution based on the conditional inference. For more details about this method see Lawless (1982), who used the conditional distributions for pivotal quantities of the parameters given the ancillary statistics as tools for estimating the parameters.

Let $Z_1=\alpha/\hat{\alpha}$, $Z_2=(\beta/\hat{\beta})^{\hat{\alpha}}$ be pivotal quantities for the parameters α and β

respectively, depending on their maximum likelihood estimators $\hat{\alpha}$ and $\hat{\beta}$ and $a_i = (x_i\hat{\beta})^{\hat{\alpha}}$ for $i = 1, 2, \cdots, n$ are the ancillary statistics. Make the change of variables from $(x_1, x_2, x_3, \cdots, x_n)$ whose pdf is (1.1) to $(\hat{\alpha}, \hat{\beta}, a_1, a_2, \cdots, a_{n-2})$. This transformation can be written as $x_i = \hat{\beta}^{-1} a_i^{1/\hat{\alpha}}$, $i = 1, 2, \cdots, n-2$, $x_{n-1} = \hat{\beta}^{-1} a_{n-1}^{1/\hat{\alpha}}$, and $x_n = \hat{\beta}^{-1} a_n^{1/\hat{\alpha}}$, where a_n and a_{n-1} can be expressed in terms of $a_1, a_2, \cdots, a_{n-2}$. The Jacobian of this transformation is independent of Z_1 and Z_2. Making further the change of variables from $(\hat{\alpha}, \hat{\beta}, a_1, a_2, \cdots, a_{n-2})$ to $(Z_1, Z_2, a_1, a_2, \cdots, a_{n-2})$, the Jacobian of this transformation is proportional to $1/Z_1 Z_2$.

Finally the conditional pdf of Z_1 and Z_2 given $A = (a_1, a_2, \cdots, a_{n-2})$ can be derived as

$$f(Z_1, Z_2 \mid A) = C Z_1^{n-1} Z_2^{-nZ_1-1} \prod_{i=1}^{n} a_i^{-Z_1} \exp(-Z_2^{-Z_1} \sum_{i=1}^{n} a_i^{-Z_1}),$$

where C is normalizing constant independent of Z_1 and Z_2.

The marginal densities of Z_1 and Z_2, conditional on A are given respectively as:

$$g_1(Z_1 \mid A) = C\Gamma(n) Z_1^{n-2} \prod_{i=1}^{n} a_i^{-Z_1} (\sum_{i=1}^{n} a_i^{-Z_1})^{-n} \tag{3.1}$$

$$g_2(Z_2 \mid A) = C \int_0^\infty Z_1^{n-1} Z_2^{-nZ_1-1} \prod_{i=1}^{n} a_i^{-Z_1} \exp(-Z_2^{-Z_1} \sum_{i=1}^{n} a_i^{-Z_1}) dz_1 \tag{3.2}$$

where

$$C^{-1} = \Gamma(n) \int_0^\infty Z_1^{n-2} \prod_{i=1}^{n} a_i^{-Z_1} (\sum_{i=1}^{n} a_i^{-Z_1})^{-n} dz_1$$

From (3.1) and (3.2) we can derive the desired probabilities for Z_1 and Z_2 and then the confidence intervals for the parameters α and β which are $(\hat{\alpha} L_{Z_1}, \hat{\alpha} U_{Z_1})$ and $(\hat{\beta} L_{Z_2}^{1/\hat{\alpha}}, \hat{\beta} U_{Z_2}^{1/\hat{\alpha}})$ respectively, where L_Z and U_Z are the lower and upper confidence interval limits for Z_1 and Z_2.

4. Simulation Study and Comparisons

The statistical performances of the proposed procedure have been compared, via Monte Carlo simulation, to the performances of the classical conditional inference in terms of the covering percentage (CP), which is defined as the fraction of times the confidence intervals cover the true value of the parameter in repeated sampling and their average lengths. The comparative results, based on 1000 Monte

Carlo simulations are given for samples of sizes $n = 20(20)100$ which have been generated for values of scale parameter $\beta = 2$ and the shape parameter $\alpha = 3$.

The simulation results in <Tables 1-2> are quite favorable to the proposed procedure. Firstly, it indicated the mean lengths of the 90% and 95% confidence intervals based on the kernel are too close for one decimal places to the ones based on the conditional inference for the parameter α, but for the parameter β the mean lengths are smaller than those for the conditional inference and getting smaller as the nominal levels and sample sizes increase. Secondly, despite the intervals based on the conditional inference for the parameter β are wider than those based on the proposed procedure, however the covering percentage based on the kernel is much greater for all sample sizes for both parameters, and it increases as the nominal level increases even for small samples. Thus the simulation results indicated a good improvement for the kernel estimates as the sample sizes increase and it can perform well and attain reasonably accurate inferences even when the number of bootstraps is extremely small up to 20 replications.

<Table 1> The Kernel and conditional mean length of intervals (MLI) for Z_1 and α, the covering percentage (CP) of the 90% and 95% confidence intervals for α.

Approaches	n	MLI (Z_1)		MLI (α)		CP (α)	
		90%	95%	90%	95%	90%	95%
Kernel	20	0.5931	0.6961	1.9908	2.3392	0.943	0.998
	40	0.4314	0.5125	1.3682	1.6249	0.949	0.987
	60	0.3490	0.4142	1.0855	1.2883	0.918	0.967
	80	0.3018	0.3586	0.9303	1.1054	0.914	0.961
	100	0.2688	0.3193	0.8267	0.9821	0.932	0.969
Conditional	20	0.5728	0.6824	1.8281	2.1781	0.859	0.946
	40	0.4055	0.4831	1.2528	1.4925	0.894	0.950
	60	0.3313	0.3948	1.0133	1.2073	0.898	0.953
	80	0.2869	0.3418	0.8727	1.0398	0.910	0.958
	100	0.2567	0.3058	0.7792	0.9285	0.916	0.951

<Table 2> The Kernel and conditional mean length of intervals (MLI) for Z_2 and β, the covering percentage (CP) of the 90% and 95% confidence intervals for β.

Approaches	n	MLI (Z_2)		MLI(β)		CP(β)	
		90%	95%	90%	95%	90%	95%
Kernel	20	1.1092	1.3332	0.6291	0.7629	0.955	0.992
	40	0.6587	0.7915	0.4109	0.4934	0.960	0.990
	60	0.5163	0.6180	0.3290	0.3935	0.938	0.979
	80	0.4366	0.5204	0.2184	0.3379	0.930	0.973
	100	0.3835	0.4574	0.2484	0.2962	0.926	0.968
Conditional	20	0.8702	1.1219	0.5447	0.6859	0.904	0.959
	40	0.5832	0.6809	0.3765	0.4405	0.901	0.946
	60	0.5216	0.6162	0.3359	0.3962	0.920	0.956
	80	0.5459	0.6139	0.3447	0.3888	0.938	0.972
	100	0.5631	0.6133	0.3503	0.3837	0.948	0.972

5. Numerical Example

Consider the data given by Dumonceaux and Antle (1973), represents the maximum flood levels (in millions of cubic feet per second) of the Susquehenna River at Harrisburg, Pennsylvenia over 20 four-year periods (1890-1969) as:

0.654, 0.613, 0.315, 0.449, 0.297, 0.402, 0.379, 0.423, 0.379, 0.324, 0.269, 0.740, 0.418, 0.412, 0.494, 0.416, 0.338, 0.392, 0.484, 0.265.

The MLE for the parameters α and β based on this data are given respectively as 4.3138 and 2.7906.

Thus for the purpose of comparison, the 90% and 95% probability intervals for the pivotal Z_1 and Z_2 are derived based on the kernel and the conditional approaches. The results in <Table 3> have been indicated the probability intervals for the pivotal quantities Z_1 and Z_2, and the corresponding confidence intervals for α and β, based on the kernel approach are shorter than the ones based on the conditional inference, which ensure the simulation results. In <Figure 1> the posterior densities of Z_1 based on the two approaches are quite identical in symmetric shape. <Figure 2> indicated the pdf for Z_2 based on the kernel approach is quite symmetric, on the contrary the ones based on the conditional approach is right skewed which ensure the increasing length of intervals for Z_2 and thus for β in the simulation results.

<Table 3>: The Lower (LL) and the Upper limits (UL) and the lengths of the 90% and 95% confidence intervals (CI) for the parameters Z_1, Z_2 and thus for α, β using the kernel and the Conditional approaches based on the flood data.

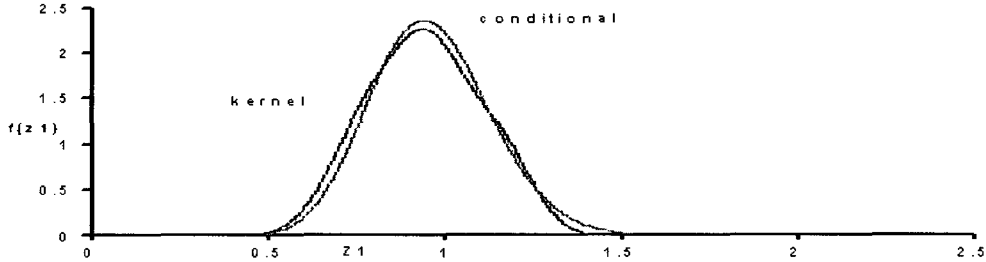
	Kernel approach				Conditional approach			
CI	90%		95%		90%		95%	
Par.	LL	UL	LL	UL	LL	UL	LL	UL
Z_1	0.6718	1.2256	0.6293	1.2669	0.6966	1.2556	0.6515	1.3176
	(0.5538)		(0.6377)		(0.5591)		(0.6661)	
α	2.8982	5.2874	2.7149	5.4660	3.0051	5.4171	2.8109	5.6845
	(2.3891)		(2.7511)		(2.4119)		(2.8736)	
Z_2	0.5572	1.2223	0.4972	1.2781	0.6576	1.5282	0.5998	1.7173
	(0.6651)		(0.7809)		(0.8706)		(1.1175)	
β	2.4368	2.9235	2.3733	2.9539	2.5322	3.0788	2.4788	3.1632
	(0.4867)		(0.5806)		(0.5466)		(0.6844)	

(The values in parentheses are the length of intervals)

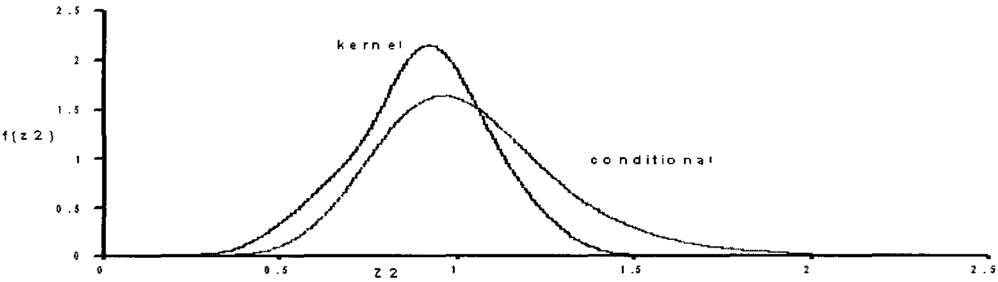
6. Conclusion

Kernel estimation technique constitutes a strong basis for statistical inference and it has a number of benefits relative to the usual conditional procedure. First, it is easy to be used and it does not need tedious work as the conditional inference. Second, it can perform well even when the number of bootstraps is extremely small up to 20 replications. Finally, it is uniquely determined on the basis of the information content in the pivotal quantities.

Thus, from the results of this paper kernel inference strengthens traditional inference statements and allows construction of alternative stronger types of inferences than the conditional inference and will encourage the statisticians for using this method.



<Figure 1> The pdf of the pivotal Z_1 based on the kernel and conditional inferences



<Figure 2> The pdf of the pivotal Z_2 based on the kernel and conditional inferences

Acknowledgments

The author would like to thank the referees and the Editor-in-Chief for their careful revision and useful comments.

References

[1] Abramson, I. (1982). On Bandwidth Variation in Kernel Estimates: A Square Root Law. *The Annals of Statistics,* Vol. 10, 1217-1223.

[2] Calabria, R. and Pulcini, G. (1990). On the maximum likelihood and least squares estimation in the inverse Weibull distribution. *Statistica Applicata,* Vol. 2, 53-66.

[3] Calabria, R. and Pulcini, G. (1992). Bayes probability intervals in a load-strength model. *Communications in Statistics-Theory & Method,* Vol. 21, 1811-1824.

[4] Dumonceaux, R. and Antle, C.E. (1973). Discrimination between the Log -normal and Weibull Distribution. *Technometrics,* Vol. 15, 923-926.

[5] Guillamon, A. Navarro, J. and Ruiz, J.M. (1998). Kernel density estimation using weighted data. *Communications in Statistics-Theory & Method*, Vol. 27, 2123-2135.

[6] Guillamon, A. Navarro, J. and Ruiz, J.M. (1999). A note on kernel estimators for positive valued random variables. Sankhya: *The Indian Journal of Statistics*. Vol. 6, Ser. A, 276-281.

[7] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions-II*. 2-nd edition. John Wiley & Sons.

[8] Jones, M.C. (1991). Kernel density estimation for length biased data. *Biometerika*, Vol. 78, 511-519.

[9] Kulczycki, P. (1999). Parameter Identification Using Bayes and Kernel Approaches. *Proceedings of the National Science Council ROC(A)*, Vol. 23, 205-213.

[10] Lawless, J.F. (1982). Statistical Models & Methods for Lifetime Data, *John Wiley & Sons*. New York.

[11] Maswadah, M. (2003). Conditional confidence interval estimation for the Inverse Weibull distribution based on censored generalized order statistics. *Journal of Statistical Computation & Simulations*, Vol. 73, 87-898.

[12] Scott D.W. (1992). *Multivariate Density Estimation*. Wiley-interscience. New York.

[13] Terrell, G.R. (1990). The maximal smoothing principle in density estimation. *Journal of the American Statistical Association*, Vol. 85, 470-477.

[Received February 2005, Accepted September 2006]