Point and Interval Estimates of Weibull Distribution to Progressively Type II Censored Data by order statistics Approach

Sujatha.V, Ravanan. R, Ramakrishnan. M

***Abstract*: In recent years, progressive censoring had a tremendous development in life testing models. If the samples do not experience the failure until the failure time period or up to the study period, then the inference on type I censoring models leads to poor statistical analysis. Instead of fixing the time failure, if number of failure components is fixed then type II censoring leads to some information results, associated with the model parameters. In both the censoring schemes none of the sample units are removed during the experiment period. But in progressive censoring scheme the number of observations and removals of the samples are fixed. In this study, point and interval estimates of the Weibull distribution for progressively type II censored data were estimated by maximum likelihood parameters, and also an exact confidence interval and region are constructed. A numerical example is presented over here to illustrate the proposed method.**

IndexTerms:Confidence-interval,maximumlikelihoodparameter, progressive type II censored samples, point and interval estimates,type II censoring

I. INTRODUCTION

In progressive type II censoring scheme, the failure components are removed along with some experimental units under test in life testing models. The basic idea in progressive type II censoring is, from a total of n units, m units are observed till the failure of the k^{th} experimental unit, where k < n is a pre-fixed integer and $\sum_{k=1}^{m} R_k + m = n$,whenever the first failure is observed, R_1 units are chosen randomly and it is removed along with that failure component , and when the second failure is observed R_2 units are chosen at randomly and it is removed with the 2^{nd} failure component i.e. from the remaining n - R_1 - 2 units and so on.

Let X (1) ≤ ... ≤ X(n) be the order statistics from a random sample of size n from a parametric distribution with the density function and with its cumulative distribution function with its own parameter, but we may not continue the experiment until the last failure since the waiting time for the final failure is unbounded (Muenz and Green, 1977).For this reason one has to pre fix the r^{th} failure, which is said to be as type-II censoring scheme.

Let X(1), ⋯ , X(n) denote the ordered values of the random sample X1, ⋯ , Xn (failure times). In Type-II plan, observations terminate after the r^{th} failure occurs (Balakrishnan and Aggarwala, 2000). So we only observe the r smallest observations in a random sample of n items. The likelihood function based on X(1), ⋯ , X(r) is given by (Arnold et al. 1992)

$$Lcen = \frac{n!}{(n-r)!}\prod_{i=1}^{r} f(x_i)[1-F(x_r)]^{n-r}$$

In Type-II censoring, the number of failure-times r is fixed whereas the endpoint X(r) is a random observation. Weibull distribution is the most important failure time distribution and it is widely used in reliability analysis. The probability density function on two parameter is f(x)=$\left\{\frac{\vartheta}{\beta}\left(\frac{x}{\beta}\right)^{\vartheta-1} \exp\left\{\frac{-x}{\beta}\right\}^{\vartheta}\right.$, for x > 0, $\beta,\vartheta > 0$, β,ϑ , are scale and shape parameters, if $0 < \beta < 1$,then the hazard function decreases, if $\beta > 1$, the hazard function increases, and if $\beta = 1$, the Weibull distribution reduces to an one parameter exponential distribution, based on one of the three rate functions for a Weibull family.[Hoyland and Rausand],various methods are there to estimate the parameters like maximum likelihood, partial likelihood, ordinary least square method and numerical methods.

II. ALGORITHM TO GENERATE A TYPE II CENSORED SAMPLES

step 1: ordered the samples in order with their failed components $X_{1:n} < X_{2:n} < X_{3:n} \ldots X_{n:n}$

step 2: $N_i = \{1, 2, \ldots n\}$; i= 1;

step 3: let k_i = min N_i and $X_{i:m:n}^{R_i} = X_{k_{i:n}}$

step4: choose randomly without a replacement sample $R_i \subseteq N_i \backslash \{k_i\}$ with $|R_i|=R_i$;

step 5: if i < m, set $N_{i+1}=Ni\backslash\{k_i\} \cup R_i$ then go to step 3 or else stop

step 6: hence,

$$X_{1:m:n}^{R_i}, X_{2:m:n}^{R_i}, X_{3:m:n}^{R_i} \ldots X_{m:m:n}^{R_i} = X_{k_{1:n}}, X_{k_{2:n}}, X_{k_{3:n}} \ldots X_{k_{m:n}}$$

A. Point Estimators

In this study, we estimated a maximum likelihood parameters for the Weibull distribution, for the progressively type II censored scheme,

Revised Manuscript Received on 30 March 2019.

* Correspondence Author
Sujatha.V* Research Scholar MTW University, Department of Mathematics, VIT University, Vellore. India.
Ravanan.R Professor, Department of Statistics, Presidency College, Chennai, India.
Ramakrishnan.M. Assistant Professor, Department of Mathematics, RKM Vivekananda college,Chennai,India

the cumulative distribution function of the Weibull distribution is

$$F(x) = 1 - \exp\left(\frac{-x}{\beta}\right)^{\vartheta}, x > 0$$

Consider $X_{1:m:n}, X_{2:m:n}, \ldots X_{m:m:n}$ are progressively type II censored sample for a two parameter Weibull distribution, with the censoring scheme $r=(r_1,r_2,\ldots r_m)$, the log- likelihood function of ϑ, β is given by,

$$L(\vartheta,\beta) = k\prod_{i=1}^{m} f(x_{i:m:n})\left[1-F(x_{i:m:n})\right]^{r_i}$$

$$k\left(\frac{J}{\beta}\right)^{m}\prod_{i=1}^{m}\left(\frac{x_{i:m:n}}{\beta}\right)^{J-1}\exp\left\{-\sum_{i=1}^{m}(r_i+1)\left(\frac{x_{i:m:n}}{\beta}\right)^{J}\right\} \quad n(n-1-r_1)$$

$(n-2-r_1-r_2)\ldots(n-m+1-r_1$

where $k = n(n-1-r_1)(n-2-r_1-r_2)\ldots(n-m+1-r_1)$

$$l(\upsilon,\beta) = \log C_1 + m\log\upsilon - m\log\beta + (\upsilon-1)\sum_{i=1}^{m}\log x_i - \frac{1}{\beta}\sum_{i=1}^{m} x_i^{\upsilon} - \frac{1}{\beta}\sum_{i=1}^{k} R_i T_i^{\upsilon}$$

$\upsilon = d(X_{1:m:n})$.

$$\beta = \frac{1}{m}\sum_{j=1}^{m}\gamma_j\left[d(X_{j:m:n}^{\beta} - X_{j-1:m:n}^{\beta}\right] = \frac{1}{m}\left\{\sum_{j=1}^{m}(R_j+1)\, X_{j:m:n}^{J}\right\}^{\frac{1}{J}} \quad ---(1)$$

and

$$\beta = \frac{1}{J} + \frac{1}{m}\sum_{j=1}^{m} X_j - \sum_{j=1}^{m}\log X_j - \frac{\sum_{j=1}^{m} R_j + 1 X_j^{J}\log X_j}{\sum_{j=1}^{m} R_j + 1 X_j^{J}} \quad ---(2)$$

This cannot be solved by analytically, hence numerical methods like Newton- Raphson method is to be employed to solve this equation.

B. Interval estimation of the parameters

Generally there are two approaches to estimate the parameter, point and interval estimation. In point estimation a numerical value of the k dimensional vectors ϑ, β was estimated. In interval estimation a k-dimensional region is determined by the parameters, specified by the level of confidence $1-\alpha$. In this study the exact confidence interval for ϑ and the exact confidence region for ϑ, β were investigated.

Let $X_1 < X_2 < \ldots X_n$ denote a type II censored from Weibull distribution, with a censoring scheme $r=(r_1,r_2,\ldots r_m)$, let $Y_i = \left(\frac{X_i}{\beta}\right)^{\vartheta}$, and $Y_1 < Y_2 < \ldots Y_m$ is also a progressively type II censored sample from exponential distribution with mean 1.

Let the transformation be

$S_1 = nY_1$

$S_2 = n - r_1 - 1(Y_2 - Y_1)$

$S_m = n - r_1 - \ldots\ r_{m-1} - m+1\ (Y_m - Y_{m-1})$

where, $S_1, S_2, \ldots S_m$ are identically independent distributed random variables which follows an exponential distribution with mean 1

hence,

$$V = 2S_1 = 2nY_1$$

has a chi-square distribution, with 2 degrees of freedom, and $U = 2\sum_{i=2}^{m} S_i = 2\left\{\sum_{i=1}^{m}(r_i+1)Y_i - nY_1\right\}$ has a chi-square distribution, with 2m-2 degrees of freedom let

$$T_1 = \frac{U}{(m-1)V} = \frac{\sum_{i=1}^{m}(r_i+1)Y_i - nY_1}{n(m-1)Y_1}$$

$$T_2 = U+V = 2\left(\sum_{i=1}^{m}(r_i+1)Y_i\right.$$

Where, U, V, T1, T2 are independent variables.

Interval estimations are calculated by finding the exact region, and the exact regions are given by the following theorems.

Theorem 1

Let $X_1 < X_2 < \ldots < X_n$ be the order statistics of progressively type II censored random variables, then the exact confidence region $100(1-\alpha)$ % of interval estimates follows the F distribution, with 2m-2,2 degrees of freedom is $\frac{\sum_{i=1}^{m}(r_i+1)x_i^{\vartheta} - nx_1^{\vartheta}}{n(m-1)x_1^{\vartheta}} = t.$

If $0 < \alpha < 1$ for an event then

$$F_{1-\alpha/2(2m-2,2)} < \frac{\sum_{i=1}^{m}(r_i+1)X_i^{\upsilon} - nX_1^{J}}{n(m-1)X_1^{J}} < F_{\alpha/2(2m-2,2)}$$

i.e

$$F_{1-\alpha/2(2m-2,2)} < \upsilon < F_{\alpha/2(2m-2,2)}$$

Theorem 2

Let X_i , i =1 to m be the order statistics, and follows a 2 parameter Weibull distribution, to a progressively type II censored sample from 1 to n, the confidence region of υ,β is given by the inequalities.

$$\left\{\begin{array}{l} f\left(X_1 \ldots X_m, F_{(\{1+\sqrt{1-\alpha}\}/2)\}(2m-2,2))}\right) < \upsilon < f\left(X_1 \ldots X_m, F_{(\{1-\sqrt{1-\alpha}\}/2)\}(2m-2,2))}\right) \\ \left\{\dfrac{2\sum_{i=1}^{m}(r_i+1)X_i^{\upsilon}}{\Psi^2\ 1-\sqrt{1-\alpha}\}/2)\}(2m)}\right\}^{\frac{1}{\upsilon}} < \beta < \left\{\dfrac{2\sum_{i=1}^{m}(r_i+1)X_i^{\upsilon}}{\Psi^2\ 1+\sqrt{1-\alpha}\}/2)\}(2m)}\right\}^{\frac{1}{\upsilon}} \end{array}\right. = 1-\alpha$$

By using these theorems one can find the point, interval and exact region for type II censored random variables.

III . DATA ANALYSIS AND NUMERICAL COMPARISONS

In this study, we examine the analysis of three real data sets, and one simulated data using progressively type II censored data from two parameter Weibull distribution, to analyse the performances of the sample based estimates and predictors of the removed units of the censored sample, The computations are performed using R($_{X64}$ 3.3.5) codes.

Example 1: Nelson (1982) reported data on times to breakdown of an insulating fluid in an accelerated test conducted at various test voltages. For illustrating the method of estimation developed in this section, let us consider the following progressively Type-II right censored sample of size m = 8 generated from the n = 19 observations recorded at 34 kilovolts in Nelson's (1982) ,and it is given by Viveros and Balakrishnan (1994)

i:	1	2	3	4	5	6	7	8
Yi:	19	0.78	0.96	1.31	2.78	4.85	6.50	7.35
Ri:	0	0	3	0	3	0	0	5

Example 2: Magne Vollan Aarset explained the time to failure of 50 devices put on life test at time 0.Based on that to illustrate the Weibull distribution, Kim and Han generated progressively type II censored data. The data and the censored scheme are given in the table.

xi	0.009	0.26	0.45	0.58	0.77	1.00	1.32	1.71
Ri	1	0	2	0	3	2	0	4

Table 2: Time to Failure of 50 devices

0.1	0.2	1	1	1	1	1	2	3	6
7	1	1	12	18	18	18	18	18	21
32	36	40	45	46	47	50	55	60	63
63	67	67	67	67	72	75	79	82	82
83	84	84	84	85	85	85	85	85	86

It is also used to illustrate the bath shaped distributions.

Table 2a: Generated progressively Type-II censored data with m=35 and n=50 from failure time data presented

xi	0.1	0.2	1	1	1	1	1	2	3	6
Ri	0	0	0	3	0	0	0	0	0	0
xi	7	11	18	18	18	18	21	32	36	45
Ri	3	0	0	0	0	0	0	3	0	0
xi	47	50	55	60	63	63	67	67	75	79
Ri	0	0	0	0	3	0	0	0	0	0
xi	82	84	84	85	86					
Ri	0	3	0	0	0					

Example 3

In this example, a progressively censored sample was generated from the Weibull lifetime data studied in Wu et al. [1] using the censoring scheme r = (1, 0, 2, 0, 3, 2, 0, 4). In the experiment, twenty devices are placed on test simultaneously and the eight observed ordered failure times, were were given below. Thus, R_i = 8 and n = 20, the available sample is x = (x1 . . . x8)

Example 4

In this example, we generate a progressive Type-II censored random sample from two parameter Weibull distribution. The following algorithm is followed to get such sample.

1. Specify the value of *n*.
2. Specify the value of *m*.
3. Specify the values of parameters β, ϑ and *p*.
4. Generate a random sample with size *m* from Weibull (β, ϑ) and sort it.
5. Generate a random number *r*1 from bino$(n - m, p)$.
6. Generate a random number *ri* from bino $(n - m - \sum_{\ell=1}^{i-1} r_\ell, p)$,for each *i*, *i* = 2, 3,... ,m-1, . .
7. Set r_m according to the following relation

$$r_m = \begin{cases} n - m - \sum_{\ell=1}^{m-1} r_\ell, & if\ n - m - \sum_{\ell=1}^{m-1} r_\ell > 0 \\ 0 & otherwise \end{cases}$$

In this sample, we assume the exact values of parameters and p are respectively 0.5, 0.5 and 0.2. In this sample,

n = 15 and *m* = 12

t1	0.278	2.009	6.352	8.286	18.325	19.332	20.333	24.727	25.717	25.877	41.47	84.676
r1	1	1	1	0	0	0	0	0	0	0	0	0

IV R CODE

R codes to find the maximum likelihood estimators of $(\hat{\vartheta}(gam\ 1), \bar{\beta}(be))$, 95% confidence limits for shape, exact confidence region, and 95% confidence interval is given below.

```
#finding the value of neu and beta
progcens=function(t,r,gam1,m,n)
  {
   neubeta=function(t,r,gam1,m)
   {
  for(i in 1:10)
 {
  gam=gam1

nr=(1/gam)+(1/m)*sum(log(t))-((sum((r+1)*(t^gam)*log(t)))
)/(sum((r+1)*(t^gam)))
   nr
   nr1=sum((r+1)*(t^gam)*log(t))
   dr1=sum((r+1)*(t^gam))

 nr2=dr1*(sum((r+1)*(t^gam)*(log(t)^2)))-nr1*sum((r+1)*
(t^gam)*log(t))
  dr=-1/(gam^2)-(nr2/(dr1*dr1))
   nr/dr
   gam1=gam-nr/dr
  }
  be=((1/m)*sum((r+1)*(t^gam1)))^((1/gam1))
  be
  gam1=round(gam1,digits=4)
  be=round(be,digits=4)
  #print(paste("MLE of Shape and Scale parameters
are",gam1,"and",be))

weii=data.frame("n"=n,"m"=m,"gamma"=gam1,"beta"=be)
  return(weii)
 }
  neubeta(t,r,gam1,m)
 #finding confidence interval for neu
 confneu=function(t,r,F,gam1,n,m)
 {
  for(i in 1:10)
  {
   gam=gam1
   fx=sum((r+1)*(t^gam))-n*(t[1]^gam)*(1+F*(m-1))

dfx=sum((r+1)*(t^gam)*log(t))-n*(t[1]^gam)*(log(t[1]))*(1
+F*(m-1))
   gam1=gam-(fx/dfx)
  }
  gam1=round(gam1,4)
  return(gam1)
 }
 #confneu(t,r,F,gam1,n,m)
```

```
F1=round(qf(0.025,2*m-2,2),4);F2=round(qf(0.975,2*m-2,2),4)
  confinte=function(t,r,F1,F2,gam1,n,m)
{
  lf=confneu(t,r,F=F1,gam1,n,m)
  uf=confneu(t,r,F=F2,gam1,n,m)
  #print(paste("95% confidence limits for shape are",lf,"and",uf))
 }
 confinte(t,r,F1,F2,gam1,n,m)
#To find F values
#To find new values of neu in joint confidence region
 F3=qf((1+sqrt(1-0.05))/2,2*m-2,2)
 F4=qf((1-sqrt(1-0.05))/2,2*m-2,2)
 confinte(t,r,F4,F3,gam1,n,m)
#To find chisquare critical values
 chi1=qchisq((1-sqrt(1-0.05))/2, 2*m)
 chi2=qchisq((1+sqrt(1-0.05))/2, 2*m)
 #limits of beta
 neu1=confneu(t,r,F=F4,gam1,n,m)
 neu2=confneu(t,r,F=F3,gam1,n,m)
 be1=((2*sum((r+1)*(t^neu1))/chi2)^(1/neu1))
 be2=((2*sum((r+1)*(t^neu2))/chi1)^(1/neu2))
 #print(paste("95% Joint confidence limits for scale are", round(be1,4),"and",round(be2,4) ))
 ret=data.frame(neubeta(t,r,gam1,m), "LCI_S"=confneu(t,r,F=F1,gam1,n,m),"UCI_S"=confneu(t,r,F=F2,gam1,n,m),"LCI_JS"=confneu(t,r,F=F4,gam1,n,m)," UCI_JS"=confneu(t,r,F=F3,gam1,n,m),"lci"=round((((2*sum((r+1)*(t^neu1))/chi2)^(1/neu1)),4),"uci"=round((((2*sum((r+1)*(t^neu2))/chi1)^(1/neu2)),4))
 return(ret)
}
```

V Results and Discussions

	Ex:1	Ex:2	Ex:3	Ex:4
n	19	50	20	15
m	8	35	8	12
gam1	0.9743	0.776	1.026	1.0256
be	9.2254	52.345	2.238	24.018
LCI_S	0.3242	0.4113	-0.943	0.3339
UCI_S	1.7692	1.2023	1.228	1.3795
LCI_JS	0.2807	0.3843	-1.191	0.2993
UCI_JS	1.964	1.3092	1.370	1.5164
lci	5.8956	25.160	0.052	3.7294
uci	11.99	74.649	3.872	45.926
m1_int_gam	1.445	0.791	2.171 9	1.0456
m2_int_gam	1.6841	0.9249	2.562 4	1.2171
m2_int_beta	6.096	49.488	3.819	42.19
per_o_cen	42	70	40	80

From Example 1 the maximum likelihood estimates of $(\hat{\beta},\hat{\vartheta})$,the scale and shape parameters are 0.9743 and 9.2254. The lower and upper confidence limits of the shape parameter are 0.3242 and 1.7692 and the joint confidence region for the shape parameter is 0.2807 and 1.9648.The lower and upper confidence interval of scale parameter are 5.8956 and 11.9923, the new confidence interval for ϑ the shape parameter is 1.445 when β is known, and the exact region of the shape parameter is 1.6841, the new confidence interval for beta is 6.0967 where gamma is fixed, and the percentage of censoring is 42. In Example 2, the maximum likelihood estimates of $(\hat{\beta},\hat{\vartheta})$,the scale and shape parameters are 0.776 and 52.3459. The lower and upper confidence limits of the shape parameter are 0.4113 and 1.2023.The joint confidence region for the shape parameter is (0.3843, 1.3092).The lower and upper confidence interval of scale parameter is (25.1609, 74.6492), the new confidence interval for ϑ the shape parameter is 0.791 when β is known, and the exact region of the shape parameter is 0.9249, the new confidence interval for beta is 49.4883, and the percentage of censoring is 70.

From Example 3, the maximum likelihood estimates of $(\hat{\beta},\hat{\vartheta})$,the scale and shape parameters are 1.0264 and 2.238. The lower and upper confidence limits of the shape parameter is (-0.9436, 1.2283).The joint confidence region for the shape parameter is (-1.1917, 1.3707).The lower and upper confidence interval of scale parameter is (0.0526, 3.8721), the new confidence interval for ϑ the shape parameter is 2.1719 when β is known, and the exact region of the shape parameter is 2.5624, the new confidence interval for beta is 3.8195, and the percentage of censoring is 70. In Example 4, the maximum likelihood estimates of $(\hat{\beta},\hat{\vartheta})$,the scale and shape parameters are 1.0256 and 24.0183. The lower and upper confidence limits of the shape parameter is (0.3339, 1.3795).The joint confidence region for the shape parameter is (0.2993, 1.5164).The lower and upper confidence interval of scale parameter is (3.7294, 45.9267), the new confidence interval for ϑ the shape parameter is 1.0456 when β is known, and the exact region of the shape parameter is 1.2171, the new confidence interval for beta is 42.1973, and the percentage of censoring is 80. When the percentage of censoring increases, it is easy to show if estimator of the shape parameter is large then the upper confidence region is large.

VI . CONCLUSIONS

The point estimators of the progressive type II censoring on two parameter Weibull distribution are estimated by using the maximum likelihood. We provide the exact confidence interval and exact confidence region for the parameters. Three real and a simulated data sets are analysed.

REFERENCES

1. Shuo-Jye Wu,(2002) Estimations of the parameters of the Weibull Distribution with progressively censored data, Journal of Japan Statist.Soc.Vol.32, 155-163.
2. Tse Sk, Xiang (2003) have explored the problem of interval estimation for parameters of Weibull-distributed data, which are Type-II progressively censored with random removals.

3. Balakrishnan.N and Aggarwala.R (2000).Progressive Censoring theory, methods and Applications.Birkhauser,Boston.
4. Balakrishnan and Erhard Cramer, The Art of Progressive Censoring, Applications to Reliability and Quality, DOI 10.1007/978-0-8176-4807-7, Springer New York .
5. Nelson.W(1982).Applied Life Data analysis,Wiley,Newyork.
6. Viveros.R and Balakrishnan.N(1994).Interval estimation of parameters of life from progressively censored data,Technometrics,36, 84-91.
7. Cacciari.M and Montanari.G.C(1987),A method to estimate the Weibull parameters,for progressively censored tests,IEEE,Transactions on Reliability,36,87-93.
8. Kim C, Han K (2009) Estimation of the scale parameter of the Rayleigh distribution under general progressive censoring. J Korean Stat Soc 38:239–246
9. Mohammad Vali Ahmadi et.al (2013), Estimating the lifetime performance index with Weibull distribution based on progressive first-failure censoring scheme, Journal of Computational and Applied Mathematics 239 , 93–102.
10. Greenwood M (1946),The statistical study of infectious diseases. J Roy Stat Soc 109: 85–110
11. A. Hoyland and M. Rausand (1994), System Reliability Theory-Models and Statistical Methods, John Wiley & Sons, Inc.