Journal of Applied Sciences

ISSN 1812-5654

Journal of Applied Sciences 6 (3): 553-558, 2006
ISSN 1812-5654

Some Useful Formulae for Monte Carlo Simulations Relating to Burr Type II, XII Distributions

A.A. Jamjoom
Department of Mathematics, Girls College of Education, Jeddah, Saudi Arabia

Abstract: In this study we develop some useful formulae relating the estimated and the actual values of the parameters of Burr type II and Burr type XII probability distribution functions. These formulae simplify the calculations for Monte Carlo Simulations executed for the estimation of some reference statistics, for example, the reliability function. The reduction of the calculation is possible the parameters of the random number generating function should assume.In some cases of the reliability function, only one Monte Carlo simulation is necessary for a given data sample amplitude. An explanation is made of the use of these formulae when estimating the reliability function.

Key words: Reliability function, confidence interval, random number, generating function, parameter estimation

INTRODUCTION

Two probability distribution functions, introduced by Burr[1], could be used for life testing are the Burr type II and the XII . They are expressed as:

$$F_T(t)= p\,[T\le t]=\frac{1}{(1+e^{-t})^r}, \qquad -\infty\langle t\langle\infty \tag{1}$$

$$F_T(t)=p[T \le t]= 1-(1+t^c)^{-k}, \qquad t\ge 0, \quad c,k\ge 1 \tag{2}$$

The hazard functions corresponding to both distributions are, respectively given by:

$$h(t)=\frac{f(t)}{1-F(t)}=\frac{re^{-t}[(1+e^{-t})^r-1]}{(1+e^{-t})^{2r+1}} \tag{3}$$

$$h(t)=\frac{ckt^{c-1}}{1+t^c} \tag{4}$$

It is clear that the hazaed rate changes with service time in both distributions (Fig. 1and 2).

A main aim in life testing is using the experimental results obtained in a specific test program to evaluate the parameters of the assumed probability distribution function , hence the actual probability of survival or some related statistic of interest can be evaluated. A useful information is the confidence level and the corresponding confidence interval of the statistic estimated. One possible way to obtain this result is to use Monte Carlo simulations. The basic use of Monte Carlo simulation in statistic can be briefly summarized in the following: given the distribution function F(x),if one generates n random numbers x_i with distribution F(x) and the number of times

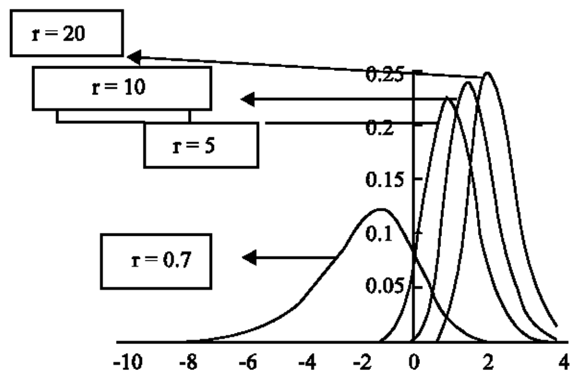

Fig. 1: Hazard function of Burr type II distribution

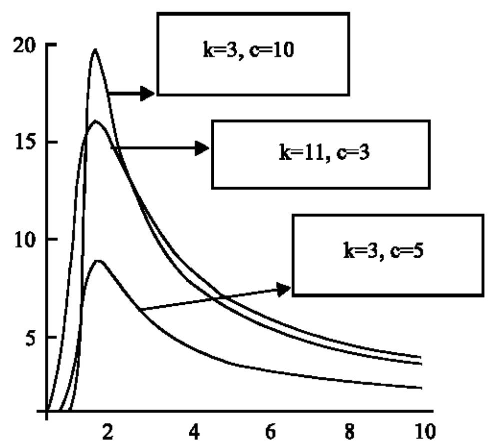

Fig. 2: Hazard function of Burr type II distribution

n_x, such that $x_i \le x$, is counted . Then the Desired estimate of F (x) is n_x/n. This same approach can be followed to obtain the confidence interval of the statistic of interest.

For the Monte Carlo simulations, to lead to a result of wide validity, it should be conducted by varying the amplitude of the data sample as well as the parameters of the random number generating function, as they are unknown a Priori. The relations to be derived in the sequel allow one to greatly reduce the amount of simulation.

PARAMETER ESTIMATION PROCEDRES

Two different procedures for parameter estimation will be considered in the sequal, both for Burr type II and XII distributions: Least Square (LS) and maximum likelihood (ML) estimators.

The LS approach: The estimated cumulative distribution is obtained by ranking the t_i values in ascending order by using

$$F_t(t) = \frac{r}{n+1}, \tag{5}$$

Where, r is the rank of the data , n is the total number of the data and t_i is the service life for the ith element of the data set in example .

By transforming Eq. 1 into the following least square fit form :

$$Z = \alpha y \tag{6}$$

Where:

$$Z_i = \ln F_T(t_i) \tag{7}$$

$$Y_i = \ln(1 + e^{-t_i}) \tag{8}$$

$$\alpha = -r \tag{9}$$

which is a straight line with slope α and intercept (0). The estimated value of the shape parameter (r) of Burr type II distribution may be obtained from the following relation Ross[2]:

$$\hat{r} = \frac{-n\sum_{i=1}^{n} Y_i Z_i + (\sum_{i=1}^{n} Y)(\sum_{i=1}^{n} Z_i)}{n\sum_{i=1}^{n} Y_i^2 - (\sum_{i=1}^{n} Y_i)^2} \tag{10}$$

In a similar way, by transforming Eq. 2 Into the following least square fit form

$$Z = \alpha y \tag{11}$$

where:

$$Z = \ln[1 - F_T(t_i)] \tag{12}$$

$$Y_i = \ln[1 + t_i^c] \tag{13}$$

$$\alpha = -k \tag{14}$$

which is a straight line with Its slope α and intercept (0). The estimated value of k, for known value of c, may be obtained from the following relation:

$$\hat{K} = \frac{-n\sum_{i=1}^{n} Y_i Z_i + (\sum_{i=1}^{n} Y_i)(\sum_{i=1}^{n} Z_i)}{n\sum_{i=1}^{n} Y_i^2 - (\sum_{i=1}^{n} Y_i)^2} \tag{15}$$

The ML Approach: consider a random sample of n independent observations with probability distribution function 2. The corresponding density function of Burr type II distribution is given by:

$$f(t) = \frac{re^{-t}}{(1+e^{-t})^{r+1}}, \quad -\infty\langle t\langle\infty, -\infty\langle r\langle\infty \tag{16}$$

The ML procedures give the value of the shape parameter r of this distribution as the result of the following equation:

$$\frac{\partial \ln L}{\partial r} = 0 \tag{17}$$

where, L is the likelihood function defined by:

$$L = \prod_{i=1}^{n} f_i(x_i) \tag{18}$$

The ln – likelihood function (LLF) for this distribution is given by

$$\ln L(r) = n \ln r - \sum_{i=1}^{n} t_i - (r+1)\sum_{i=1}^{n} \ln(1 + e^{-t_i}) \tag{19}$$

Then using Eq. (17) one can obtain the estimate of the shape parameter r of Burr type II distribution as:

$$\hat{r} = \frac{n}{\sum_{i=1}^{n} \ln(1 + e^{-t_i})} \tag{20}$$

For Burr type XII the probability density function corresponding to 2 is given by:

$$f(t) = \frac{ckt^{c-1}}{(1+t^c)^{k+1}}, \quad t \geq 0,\ c,k \geq 1 \tag{21}$$

The ML procedure given the values of the two shapes parameters c and k of this distribution as the result of the following system of equations:

$$\frac{\partial \ln L}{\partial c} = 0 \qquad (22)$$

$$\frac{\partial \ln L}{\partial k} = 0 \qquad (23)$$

Where, L is the likelihood function defined by:

$$L = \prod_{i=1}^{n} f(x_i)$$

Wingo[3].

The ln – likelihood function (LLF) for this distribution is given by

$$\ln L(c,k) = n\ln c + n\ln k + (c-1)\sum_{i=1}^{n} \ln t_i - (k+1)\sum_{i=1}^{n} \ln(1+t_i^c) \qquad (24)$$

Then

$$\frac{\partial \ln L(c,k)}{\partial k} = \frac{n}{k} - \sum_{i=1}^{n} \ln(1+t_i^c) = 0 \qquad (25)$$

from which we get

$$\hat{k} = \frac{n}{\sum_{i=1}^{n} \ln(1+t_i^c)} \qquad (26)$$

Similarly

$$\frac{\partial \ln L(c,k)}{\partial c} = \frac{n}{c} + \sum_{i=1}^{n} \ln t_i - (k+1)\sum_{i=1}^{n} \frac{t_i^c \ln t_i}{(1+t_i^c)} = 0 \qquad (27)$$

which can be written as

$$\frac{n}{\hat{c}} + \sum_{i=1}^{n} \ln t_i = (\hat{k}+1)\sum_{i=1}^{n} \frac{t_i^{\hat{c}} \ln t_i}{1+t_i^{\hat{c}}} \qquad (28)$$

where the carets refer to the values of the estimates of the shape parameters. The Estimate of c is obtained from Eq. 28 and there after the estimate of k is obtained from Eq. 27

INTERESTING RELATIONS FOR THE PARAMETER ESTIMATION WITH RELATION TO MONTE CARLO SIMULATIONS

The confidence intervals of some statistics of interest may be obtained by means of Monte Carlo simulations. Sets of random numbers with uniform distribution are generated and then, the values of the parameters of the distribution (r for Burr type II distribution, c and k for Burr type XII distribution), being fixed, they are transformed into random numbers having the distribution of interest. On the basis of the generated random number data set, the statistic of interest is then estimated. The value obtained is used to build an approximate probability distribution function for such a statistic and then to evaluate the confidence interval. To maintain a certain generality of the analyses, the simulations must be run with the parameters of the probability distribution of interest taking values in a certain bounded region, where they are likely to lie. It is clear that the definition of the amplitude of the region as well as the number of points in it considered for the simulation has a great influence on the calculation efforts and on their amplitude. With the aim of limiting the number of simulations without losing the generality of the analyses, the study of the dependence of the estimated values of the parameters on the parameter's values of the actual distribution is of certain interest.

The LS and ML estimations with reference to the Monte Carlo simulations for Burr type II and XII distributions: According to the calculation reported in the Appendix A for LS estimation, it follows
that:

$$\hat{r} = r\,k \qquad (29)$$

while from the calculation reported in the Appendix B for ML estimation, it is

$$\hat{r} = \frac{n}{\sum_{i=1}^{n} \ln\left(\frac{1-2p_i^{\frac{1}{r}}}{1-p_i^{\frac{1}{r}}}\right)} \qquad (30)$$

which depend only on the set of random numbers generated by uniform distribution and not on the actual Burr type II distribution function.

From the calculations reported in the Appendix C for LS estimation for Burr type XII
distribution, it comes out that for known value of the shape parameter c , it comes out that

$$\hat{k} = k\,\Phi \qquad (31)$$

while from the calculations reported in the Appendix D for ML estimation, it is:

$$\hat{k} = k\,\Phi'$$

$$\frac{n}{\hat{c}} + \frac{1}{c}\sum_{i=1}^{n} \ln\left[(1-p_i)^{-\frac{1}{k}} - 1\right] \qquad (32)$$

$$= \left(\hat{k}+1\right)\sum_{i=1}^{n}\left[(1-p_i)^{-\frac{1}{k}} - 1\right]^{\frac{\hat{c}}{c}} \bullet \ln\left[(1-p_i)^{-\frac{1}{k}} - 1\right]^{\frac{1}{c}} \qquad (33)$$

Where, Φ and Φ' and Eq. 33 depend only on the set of random numbers generated with uniform distribution and not on the actual Burr type XII distribution function.

Some applications of the derived relations: The formulae obtained in the previous sections will allow us to simplify the Monte Carlo simulation for some statistics of interest. Let us now consider the case of the evaluation of the reliability function, where one has to estimate the following functions,
for Bur type II and XII distributions:

$$R(t) = 1-\left(1-e^{-t}\right)^{-r} \tag{34}$$

$$R(t) = \left(1+t^{c}\right)^{-k} \tag{35}$$

whose estimated expression is given respectively by

$$\hat{R}(t) = 1-\left(1-e^{-t}\right)^{-\hat{r}} \tag{36}$$

$$\hat{R}(t) = \left(1+t^{c}\right)^{-\hat{k}} \tag{37}$$

For the case of LS estimation for the Burr type II, let us substitute Eq. 29 into Eq. 36 and obtaining the service time t from Eq. 34 it comes out:

$$\hat{R}(t) = 1-\left(1-R(t)\right)^{k} \tag{38}$$

For the case of ML estimation for Burr type II, similarly as above, let us substitute Eq. 30 into Eq. 36 and obtaining the service time t from Eq. 34, it comes out

$$\hat{R}(t)=1-\left[1-R(t)\right]^{l} \tag{39}$$

where

$$l = \frac{n}{r\sum_{i=1}^{n}\ln\frac{(1-2p_i^{\frac{1}{r}})}{(1-p_i^{\frac{1}{r}})}}$$

For the case of LS estimation for the Burr type XII, let us substitute Eq. 31 into Eq. 37 and obtaining the service time t from Eq. 35

$$\hat{R}(t) = \left(R(t)\right)^{\varphi} \tag{40}$$

For the case of ML estimation for Burr type XII, similarly as above, let us substitute Eq. 32 into 37 and obtaining the service time t from Eq. 35, it comes out:

$$\hat{R}(t)=\left(R(t)\right)^{\Phi'} \tag{41}$$

It follows from Eqs. 38 through 42 that the actual reliability function dependsonly on its estimated value and not on the actual value of its parameters. This means that if one is interested in estimating the confidence intervals of the reliability functionof Burr type II or of XII distribution, it is sufficient to run only one Monte Carlo simulation for a given data sample amplitude using any value for the parameters of the actual probability distribution function used as random number generator in thesimulations. This greatly reduces the amount of calculations involved in the analyses. Using the same properties, i.e. the results expressed by Eq. 29 through 33, also other statistics can be studied reducing the amount of calculations needed. Consider for example the following case: One has the results of some experiments, reporting the life to failure of some mechanical component. On the basis of this sample, one would like to estimate, which is the value of the shape parameters of the underlying Burr type II or XII distribution. To assess this, one would need to run some Monte Carlo simulations varying the shape parameters of the random number generating function for the given data set amplitude and evaluate which is the confidence of the estimated value of the shape parameter. Assuming to know which is the interval where the shape parameters would lie, it would be consequently necessary to run a certain number of simulations. Assuming to discretize these intervals in 10 values for the shape parameter (c) and in 5 values for the shape parameter (k), it comes out that for a given data set amplitude 50 Monte Carlo simulations have to be run . But because of the results reported in Eqs. 29 and 31 the actual value of the shape parameter depends only on its estimated value and consequently only one Monte Carlo simulation is needed, thus strongly reducing the amount of calculations required from 50 runs to 1 run in this case.

CONCLUSIONS

In this study some useful formulae are presented for reducing the amount of calculations in numerical simulation analyses for the evaluation of the estimation of the parameters and of other statistics for the Burr type II and XII probability distribution functions. The results obtained have been shown to be very useful for simulation studies for the reliability function of both distributions.

APPENDIX A

The LS estimation with reference to the Monte Carlo simulations for Burr Type II distribution: A set of random numbers is generated having uniform distribution in the interval [0,1]. Such a set, indicated as $\{p_i\}$, corresponds to the values of the probabilities. The related set of lifetime values, $\{t_i\}$, is obtained by inversion of Eq. (1)

$$t_i = \frac{1}{r}\ln p_i - \ln(1 - p_i^{\frac{1}{r}}) \tag{A1}$$

To obtain the values of the estimate of the shape parameter r, one has to apply Eq. 10 by using Eqs. 6-9 and substituting the t_i term with the expression obtained in Eq. (A1). The results are as follows:

$$\begin{aligned} -n\sum Y_i Z_i &= -n\sum_{i=1}^{n}\ln(1+e^{-t_i})(Z_i) \\ &= -n\sum_{i=1}^{n}\ln(1+e^{-(\frac{1}{r}\ln p_i - \ln(1-p_i^{\frac{1}{r}}))}).(Z_i) \\ &= -n\sum_{i=1}^{n}\ln(1+p_i^{\frac{-1}{r}}(1-p_i^{\frac{1}{r}})).(Z_i) \\ &= -n\sum_{i=1}^{n}(\ln p_i^{\frac{-1}{r}})(Z_i) \\ &= \frac{n}{r}\sum_{i=1}^{n}Z_i \ln p_i \ , \end{aligned} \tag{A2}$$

We also note that

$$\begin{aligned} (\sum_{i=1}^{n}Y_i)\ (\sum_{i=1}^{n}Z_i) &= \left[n\sum_{i=1}^{n}\ln(1+e^{-t_i})\right]\left[\sum_{i=1}^{n}Z_i\right] \\ &= -\frac{1}{r}\left[\sum_{i=1}^{n}\ln p_i\right]\left[\sum_{i=1}^{n}Z_i\right] \end{aligned} \tag{A3}$$

and also that

$$\begin{aligned} n\sum_{i=1}^{n}Y_i^2 - \left(\sum_{i=1}^{n}Y_i\right)^2 &= n\sum_{i=1}^{n}(\ln(1+e^{-t_i}))^2 - (\sum_{i=1}^{n}\ln(1+e^{-t_i}))^2 \\ &= n\ (\sum_{i=1}^{n}(-\frac{1}{r}\ln p_i)^2 - (\sum_{i=1}^{n}-\frac{1}{r}\ln p_i)^2 \\ &= \frac{n}{r^2}\sum_{i=1}^{n}(\ln p_i)^2 - \frac{1}{r^2}(\sum_{i=1}^{n}\ln p_i)^2 \end{aligned} \tag{A4}$$

Substituting A2, A3 and A4 into 10

$$\hat{r} = r\frac{\left[n\sum_{i=1}^{n}(\ln p_i).Z_i\right] - \left[\sum_{i=1}^{n}\ln p_i\right]\left[\sum_{i=1}^{n}Z_i\right]}{\sum_{i=1}^{n}(\ln p_i)^2 - \left(\sum_{i=1}^{n}\ln p_i\right)^2} \tag{A5}$$

$$\hat{r} = rk \tag{A6}$$

where:

$$k = \frac{n\sum_{i=1}^{n}Z_i(\ln p_i) - \left[\sum_{i=1}^{n}\ln p_i\right]\left[\sum_{i=1}^{n}Z_i\right]}{\sum_{i=1}^{n}(\ln p_i)^2 - \left[\sum_{i=1}^{n}\ln p_i\right]^2} \tag{A7}$$

where, k depends only on the uniform distributed random data set generated (the $\{p_i\}$ sample), i.e. it is not dependent on the Burr type Ð generating distribution function.

APPENDIX B

The ML estimation with reference to the Monte Carlo simulations for Burr Type II distribution: Let us consider now the term of Eq. 20 which give the shape parameter estimate, and let us substitute the expression of t_i obtained by inversion of Eq. 1 represented in Eq. A1; we have:

$$\hat{r} = \frac{n}{\sum_{i=1}^{n}\ln\left(1+e^{-\ln\frac{p_i^{\frac{1}{r}}}{1-p_i^{\frac{1}{r}}}}\right)}$$

$$= \frac{n}{\sum_{i=1}^{n}\ln(1-\frac{p_i^{\frac{1}{r}}}{1-p_i^{\frac{1}{r}}})} \tag{B1}$$

$$\hat{r} = \frac{n}{\sum_{i=1}^{n}\ln\left(1-\frac{1-2p_i^{\frac{1}{r}}}{1-p_i^{\frac{1}{r}}}\right)} \tag{B2}$$

which depends only on the uniform distributed data set.

APPENDIX C

The LS estimation with reference to the Monte Carlo simulations for Burr type II distribution: A set of random numbers is generated having uniform distribution in the interval [0,1]. Such a set, indicated by $\{p_i\}$, corresponds to the values of the probabilities. The related set of lifetime values $\{t_i\}$, is obtained by inversion of Eq. 2 as:

$$F_T(t) = p[T \le t] = 1-(1+t_i^c)^{-k} = p_i$$

$$t_i = \left[(1-p_i)^{-\frac{1}{k}} - 1\right]^{\frac{1}{c}} \quad \text{(C1)}$$

By substituting the value of t_i obtained in Eq. C 1 into Eq. 13 and then substituting t he result into Eq. 15, it comes out:

$$Y_i = -\frac{1}{k}\ln(1-p_i) \quad \text{(C2)}$$

and

$$\hat{k} = \frac{-n\sum_{i=1}^{n} Y_i Z_i + (\sum_{i=1}^{n} Y_i)(\sum_{i=1}^{n} Z_i)}{n\sum_{i=1}^{n} Y_i^2 - (\sum_{i=1}^{n} Y_i)^2}$$

$$\hat{k} = k\,\frac{n\sum_{i=1}^{n}\ln(1-p_i).Z_i - (\sum_{i=1}^{n}\ln(1-p_i))(\sum_{i=1}^{n} Z_i)}{n\sum_{i=1}^{n}(\ln(1-p_i))^2 - (\sum_{i=1}^{n}\ln(1-p_i))^2}$$

$$\hat{k} = k\,\phi \quad \text{(C3)}$$

where

$$\phi = \frac{n\sum_{i=1}^{n}(\ln(1-p_i)).Z_i - (\sum_{i=1}^{n}\ln(1-p_i))(\sum_{i=1}^{n} Z_i)}{n\sum_{i=1}^{n}(\ln(1-p_i))^2 - (\sum_{i=1}^{n}\ln(1-p_i))^2} \quad \text{(C4)}$$

where ϕ depends only on the set of random numbers generated with uniform distribution and not on the actual Burr type XII distribution function.

APPENDIX D

The ML estimation with reference to the Monte Carlo simulations for Burr type XII Distribution:
substituting Eq. (C1) into Eq.(28) we get

$$\frac{n}{\hat{c}} + \sum_{i=1}^{n}\ln t_i = (\hat{k}+1)\sum_{i=1}^{n}\frac{t_i^{\hat{c}}\ln t_i}{1+t_i^{\hat{c}}}$$

$$\frac{n}{\hat{c}} + \sum_{i=1}^{n}\ln\left[(1-p_i)^{-\frac{1}{k}} - 1\right]^{\frac{1}{c}}$$

$$= (\hat{k}+1)\sum_{i=1}^{n}\left[(1-p_i)^{-\frac{1}{k}} - 1\right]^{\frac{\hat{c}}{c}}\ln\left[(1-p_i)^{-\frac{1}{k}} - 1\right]^{\frac{1}{c}}$$

$$\frac{n}{\hat{c}} + \frac{1}{c}\sum_{i=1}^{n}\ln\left[(1-p_i)^{-\frac{1}{k}} - 1\right] \quad \text{(D1)}$$

$$= (\hat{k}+1)\sum_{i=1}^{n}\left[(1-p_i)^{-\frac{1}{k}} - 1\right]^{\frac{\hat{c}}{c}}\ln\left[(1-p_i)^{-\frac{1}{k}} - 1\right]^{\frac{1}{c}}$$

which is an implicit equation in to be solved recursively and it depends only on the set of random numbers with uniform distribution considered.

Let us now substitute Eq. C1 into Eq. 26 we then have .

$$\hat{k} = \frac{n}{-\frac{1}{k}\sum_{i=1}^{n}\ln(1-p_i)}$$

$$= k\,\frac{-n}{\sum_{i=1}^{n}\ln(1-p_i)}$$

$$\hat{k} = k\,\Phi' \quad \text{(D2)}$$

where:

$$\Phi' = \frac{-n}{\sum_{i=1}^{n}\ln(1-p_i)} \quad \text{(D3)}$$

depends only on the uniform distributed data set.

REFERENCES

1. Burr, I.W., 1942. Cumulative frequency functions. Ann. Math. Statist. 13: 215-232.
2. Ross, M.S., 1987. Introduction to Probability and Statistics for Engineers and Scientists. John Wiley and Sons, New York, pp: 380-399.
3 Wingo, D.R., 1983. Maximum likelihood methods for fitting the Burr type XII distribution to life test data. Biom. J., 25: 77-84.