On Extension of Weibull Model with an Application

T.A. Raja, R. K. Agnihotri* and Bilal Ahmad**
DOS, Mirgund, S.K. University of Agricultural Sciences and Technology Kashmir (J&K)
* St. John's College Agra (U.P)
** Division of Agri. Statistics, SKUAST-K

ABSTRACT

This paper Studies a new Statistical model, which yields well known weibull and exponential probability models as a particular case. The estimation of the Parameters has been found by the method of maximum Likelihood. The application part has been exploited by a numerical illustration of average maximum temperature (degrees in Celsius) and relative humidity (percentage) averaged over ten years(1997-2006) for district Srinager of Kashmir valley with the help of Software package.

Keywords: Statistical model, estimation, probability density function, temperature, relative humidity

INTRODUCTION

In the literature, we come across different models e.g., Linear Models, Non-Linear Models, Regression analysis Models, Generalized Linear Models, Generalized Additive Models, Analysis of Variance Models, Stochastic Models, Simulation Models, Operation Research Models, Catalytic Models, Markov-based Models etc. Out of large number of methods and tools developed so for analyzing the data the Statistical modeling is the latest innovations Gilchrist warren (1984).Models are considered as the back bone of modern Statistics and data analysis. There exists a large number of distributions of continuous type [e.g., Hogg & Craig; Johnson & Kotz and Lawless J.F.] Since a discrete as well as distribution stands on some stimulated assumptions and any variation in these assumptions leads to a different situation. It is natural to study a modification and revision of a distribution depending upon the nature of change in the situation or violation of assumptions which give rise to a new class of distributions. In general, it is possible to improve the fit of a distribution by the incorporation of extra parameters or variation in the parameters. Mudholkar *et. al.,* (1993,1995) and Bilal *et al.,* (2005) proposed models with the probability density function of Weibull type. In this paper, we also propose a model as.

PROPOSED STATISTICAL MODEL

THEOREM 1. For any real numbers $\alpha,\ p, \theta$ *and* β, the function

$$f_X(x)=\frac{p}{\theta}\left(\frac{\alpha x+\beta}{\theta}\right)^{p-1}\alpha\exp\left[-\left\{\frac{\alpha x+\beta}{\theta}\right\}^{p}\right]; \quad \ldots\ldots\ldots\ldots \quad (1.1)$$

$$p,\theta,\alpha > 0 \text{ and } \beta \geq 0$$

$= 0$, elsewhere

is the probability density function of of the random variable X of continuous type. Where $p, \theta, \beta, and\, \alpha$ are the parameters of the model.

PROOF OF THEOREM 1. To prove $f_X(x)$ is a probability density function, following conditions are to be satisfied:

(i) $f_X(x) \geq 0$ and (ii) $\int_{-\infty}^{\infty} f_X(x)dx = 1$

Since p, θ > 0; α ≥ 0, β > 0, therefore it follows that $f_X(x) > 0\ \forall\, x$. Also $f_X(x) = 0$ elsewhere. Hence $f_X(x) \geq 0\ \forall\, x$, which proves condition (i).

clearly . $f_X(x) \geq 0$

Also to establish condition (ii), we have

$$\int_{-\infty}^{\infty} f_X(x)dx = \int_{-\infty}^{\infty}\frac{p}{\theta}\left(\frac{\alpha x+\beta}{\theta}\right)^{p-1}\alpha\exp\left[-\left\{\frac{\alpha x+\beta}{\theta}\right\}^{p}\right]dx \quad \ldots\ldots \quad (1.2)$$

we suppose $\left(\frac{\alpha x+\beta}{\theta}\right)^{p} = y.$

$$\frac{dy}{dx} = p\left(\frac{\alpha x+\beta}{\theta}\right)^{p-1}.\frac{\alpha}{\theta}$$

$$\frac{dy}{p\alpha}.\theta = \left(\frac{\alpha x+\beta}{\theta}\right)^{p-1}.dx$$

It gives on differentiation, $\frac{px}{\theta}\left(\frac{\alpha x+\beta}{\theta}\right)^{p-1}dx = \frac{\theta}{\alpha p}\,dy.$

Hence from (1.2), it follows that

$$\int_{-\infty}^{\infty} f_X(x)dx = \int_0^{\infty} e^{-y} dy = 1.$$

Thus, we have proved that

(i) $f_X(x) \geq 0$ for all x and

(ii) $\int_{-\infty}^{\infty} f_X(x)dx = 1.$

This completes the proof of Theorem 1.

ESTIMATION OF THE PARAMETERS

There are different methods for estimating the parameters available in the literature e.g., the method of moments, the method of maximum likelihood, the method of Bayesian estimation, the method of least square, entropy method etc. It is interesting to note that researchers are trying to find out which estimation method is preferable for the parameter estimation of a particular probability distribution in order to get reliable estimates of the parameters.

The Maximum Likelihood is the most efficient and most preferred method for estimating the parameters but, sometimes, it involves complicated forms of Maximum Likelihood equations which are difficult to solve for Maximum Likelihood Estimators. In such case, some other efficient estimators are to be find out. In this section, we will study the Maximum Likelihood Method for estimating the parameters of the model (1.2) as follows:

Consider a random sample $x_1, x_2, \ldots, x_n$ of size n from the population (1.2). Then, the likelihood Function L is given by

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

$$= \prod_{i=1}^{n} \left[\frac{p}{\theta} \left[\frac{\alpha x+\beta}{\theta}\right]^{p-1} .\alpha. \exp\left[- \left\{ \frac{(\alpha x_i + \beta)}{\theta} \right\}^{p} \right] \right]$$

Therefore

$$\text{Log L} = \sum_{i=1}^{n} \log p - \sum_{i=1}^{n} \log\theta + \sum_{i=1}^{n} (p-1)[\log(\alpha x_i + \beta) - \log\theta]$$

$$+ \sum_{i=1}^{n} \log\alpha - \left\{ \frac{(\alpha x_i + \beta)}{\theta} \right\}^{p}$$

The likelihood equation for the estimation of the parameters , p, θ, α *and* β, will be as

$$\frac{\partial}{\partial p} LogL = 0 \ldots \ldots \qquad (2.1)$$

$$\sum_{i=1}^{n} \frac{1}{p} + \sum_{i=1}^{n} ((Log(\alpha xi + \beta) - Log\theta) = 0$$

$$\sum_{i=1}^{n} \frac{1}{p} + \sum_{i=1}^{n} Log\left\{\frac{\alpha xi + \beta}{\theta}\right\} = 0 \ldots \ldots \qquad (2.2)$$

$$\frac{\partial}{\partial \theta} LogL = 0 \ldots \ldots \ldots . \qquad (2.3)$$

$$-\sum_{i=1}^{n} \frac{1}{\theta} - \sum_{i=1}^{n} \frac{p}{\theta} + \sum_{i=1}^{n} \frac{1}{\theta} - p(\theta)^{-p-1} (\alpha xi + \beta)^{p} = 0..$$

$$-\sum_{i=}^{n} \frac{p}{\theta} - \frac{p}{\theta^{p+1}} (\alpha xi + \beta) = 0 \ldots \ldots \ldots \qquad (2.4)$$

$$\frac{\partial}{\partial \beta} LogL = 0 \ldots \ldots \ldots \ldots \ldots . \qquad (2.5)$$

$$\sum_{i=1}^{n} p \frac{1}{(\alpha xi + \beta)} . \frac{1}{\beta} - \sum_{i=1}^{n} \frac{1}{(\alpha xi + \beta)} \frac{1}{\beta} - p\left\{\frac{\alpha xi + \beta}{\theta}\right\}^{p-1} \frac{1}{\theta} = 0 \ldots \qquad (2.6)$$

$$\frac{\partial}{\partial \alpha} LogL = 0 \ldots \ldots \ldots . \qquad (2.7)$$

$$\sum_{i=1}^{n} p . \frac{1}{(\alpha xi + \beta)} . x - \sum_{i=1}^{n} \frac{1}{(\alpha xi + \beta)} . x + \sum_{i=1}^{n} \frac{1}{\alpha} - p\left\{\frac{(\alpha xi + \beta}{\theta}\right\}^{p-1} . \frac{xi}{\theta} = 0 .. (2.8)$$

Estimation of the parameters $p, \theta \ \beta, and\, \alpha$ are obtained for solving equations (2.2), (2.4), (2.6) and (2.8).(e.g Rao,C.R)

DERIVATION OF VARIOUS DISTRIBUTIONS

In fact, a number of well-known distributions follow from the model (1.1) for a suitable choice of the parameters $p, \theta \ \beta, and\, \alpha$. Here we mention only a few

REMARK 3.1. Taking $\alpha = 1$, $p = 1$ and $\beta = 0$ in (1.1), we have

$$f(x) = \frac{p}{\theta}\left(\frac{x}{\theta}\right)^{p-1} \exp\left[-\left\{\frac{x}{\theta}\right\}^{p}\right],$$

$= 0$, elsewhere

which is the p.d.f of well-known Weibull Distribution with parameters $p(>0)$ and $\theta(>0)$.

REMARK 3.2. Taking $\alpha = 1$, $p = 1$ and $\beta = 1$ in (1.1), we have

$$f(x) = \frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right)$$

$= 0$, elsewhere

which is the p.d.f of well-known Exponential Distribution with parameter $\theta(>0)$.

Similarly as above, many more distributions can be shown as a particular case of the proposed statistical model (1.1) for a suitable of choice of the parameters

APPLICATION

The fitting of two-parameter Weibull distribution and exponential distribution on maximum temperature(degree in Celsius) and relative humidity (%) averaged over ten years(1997-1998 to 2006-2007) for district Srinager in Kashmir valley has been exploited (Table 1.)

Table 1. Depicting the average maximum temperature (degrees in Celsius) at Srinager

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1997	3.2	10.8	14.2	19.4	22.5	27.4	31.9	29.1	27.8	19.9	13.7	8.4
1998	5.8	10.9	16.2	21.2	25.4	28.9	30.6	29.6	27.6	23.1	16.4	9.5
1999	5.9	10.7	15.5	24.7	26.0	30.3	31.6	29.9	29.7	24.7	15.4	12.6
2000	7.6	10.2	15.2	23.1	29.1	30.1	30.0	29.6	27.1	25.6	17.2	11.0
2001	11.7	13.7	18.0	21.7	28.4	29.6	30.1	30.2	26.9	24.4	15.5	10.6
2002	9.3	9.5	16.8	20.6	26.5	28.9	30.5	29.8	25.1	23.3	18.5	10.0
2003	11.2	10.3	13.7	21.0	22.2	30.1	30.9	28.6	26.9	23.0	15.4	10.0
2004	7.1	13.0	21.7	20.7	25.4	27.8	29.4	29.3	29	20.6	17.9	9.7
2005	7.5	6.5	14.7	20.7	21.8	29.3	28.9	30.4	29.3	22.7	15.8	9.9
2006	4.3	13.4	16.0	21.1	28.2	27.6	30.9	28.7	25.9	22.9	15.0	8.4
G.M	7.50	10.90	16.20	21.42	25.55	29.00	30.48	29.52	27.53	23.02	16.08	10.06

(Source:- Regional Metrological Centre, Srinager; Digest of Statistics(2006-2007).

Table 2 Depicting the average relative humidity in (percentage) at 1730 hours at Srinagar

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1997	67	60	50	46	46	45	50	58	58	62	61	70
1998	65	58	51	51	45	45	52	59	54	65	60	67
1999	58	57	46	48	45	46	53	60	55	62	62	66
2000	57	59	48	49	48	47	50	56	59	64	64	69
2001	58	49	44	47	43	52	56	55	52	55	61	69
2002	62	60	47	51	44	48	51	60	58	80	53	65
2003	56	62	57	50	51	47	50	60	62	63	62	69
2004	70	52	34	56	46	50	51	55	49	63	73	74
2005	67	78	61	41	58	44	60	56	50	56	56	64
2006	79	61	52	42	44	46	54	61	60	61	68	75
G.M	63.90	59.60	49.00	48.10	47.00	47.00	52.60	57.80	55.70	63.00	62.00	68.85

(Source:- Regional Metrological Centre, Srinager; Digest of Statistics(2006-2007))

Using Minitab (14 version, http://www.minitab.com) for the analysis of the data.
Distribution Analysis: Maximum temperature

Distribution: Weibull
Estimation Method: Maximum Likelihood

Table 3 depicting the parameter estimates for maximums temperatures (^{0}C)

Parameter	Estimate	Standard Error	95.0% Normal CI Lower	Upper
Shape	1.98815	0.223895	2.48585	3.36687
Scale	24.0888	0.764928	21.6979	24.6984

Log-Likelihood = -419.300
Goodness-of-Fit
Anderson-Darling (adjusted) = 3.856

Table 4 Characteristics of distribution for maximum temperature (^{0}C)

Standard 95.0% Normal CI
Estimate Error Lower Upper

Mean(MTTF) 21.3506	0.704229	19.3051		22.0677
Standard Deviation 11.2203	0.500560 6.82823		8.79564	
Median 20.0334	0.757707 18.9627		21.9355	
First Quartile(Q1) 15.0491	0.804293	13.5525		16.7110
Third Quartile(Q3) 30.5668	0.818247 24.3614		27.5709	
Interquartile Range(IQR)15.5177	0.709797	9.56161		12.3516

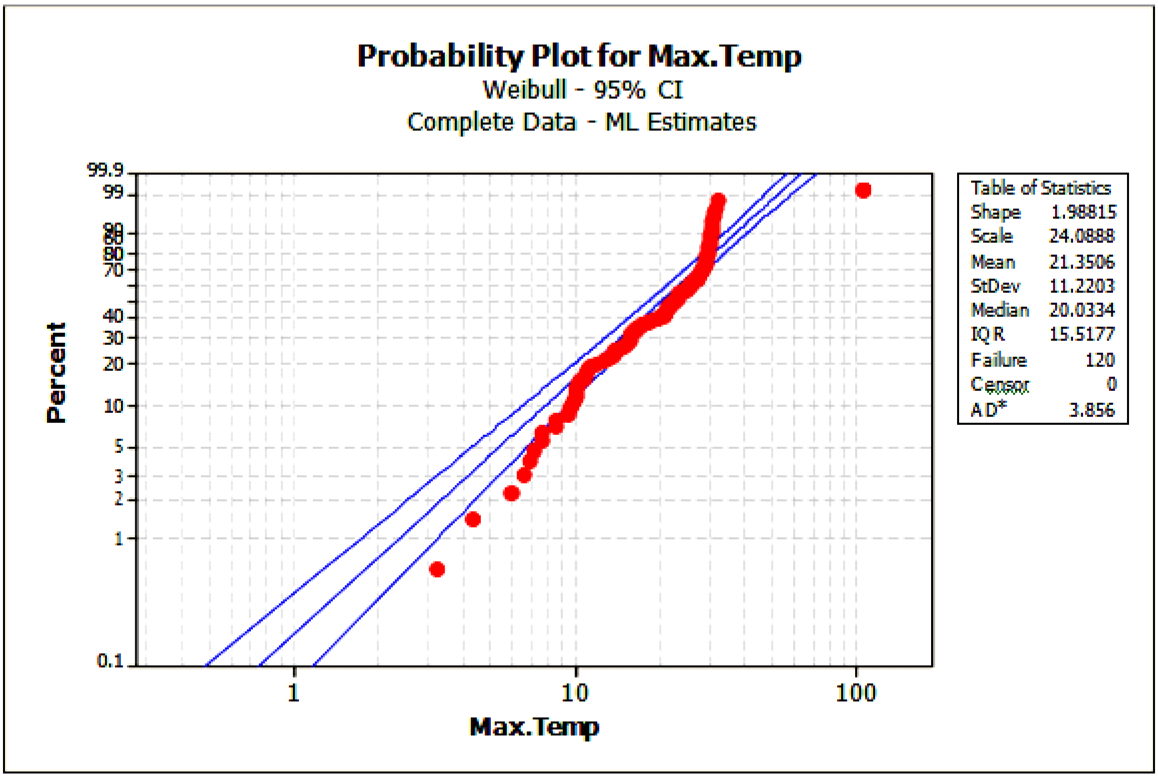

Fig. 1 Probability plot for maximum temperature (^{0}C)

Estimation Method: Maximum Likelihood

Table 5 Parameter estimates for relative humidity

		Standard	95.0% Normal CI	
Parameter	**Estimate**	**Error**	**Lower**	**Upper**
Shape	6.67887	0.444077	5.86282	7.60850
Scale	59.9703	0.869169	58.2907	61.6982

Log-Likelihood = -436.011
Goodness-of-Fit
Anderson-Darling (adjusted) = 1.559
Characteristics of Distribution

Table 6 Characteristics and distribution of relative humidity (%)

	Standard		95.0% Normal CI	
	Estimate	**Error**	**Lower**	**Upper**
Mean(MTTF)	55.9602	0.896360	54.2306	57.7449
Standard Deviation	9.82113	0.539879	8.81800	10.9384
Median	6.7680	0.912800	55.0068	58.5855
First Quartile(Q1)	49.7647	1.09424	47.6656	51.9562
Third Quartile(Q3)	62.9761	0.866413	61.3006	64.6973
Interquartile Range(IQR)	13.2114	0.779537	11.7685	14.8311

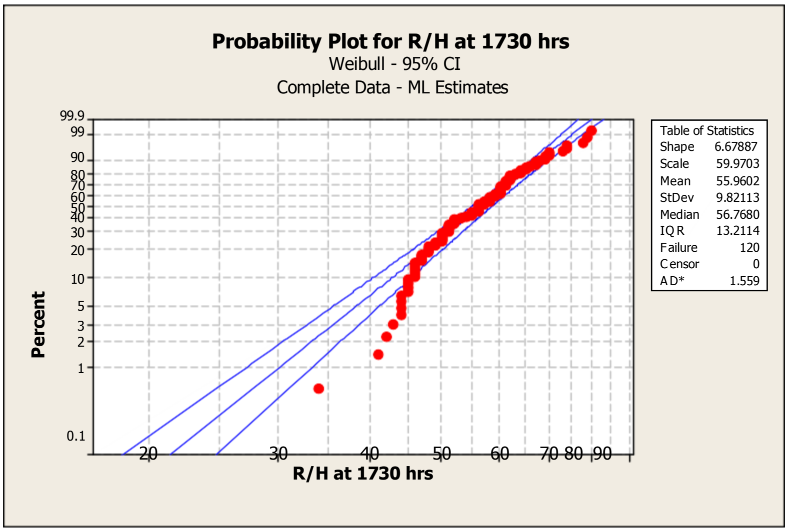

Fig 2 Probability plot for relative humidity (%)

Estimation Method: Maximum Likelihood
Distribution: Exponential

Parameter Estimates

Parameter	Estimate	Standard Error	95.0% Normal CI Lower	Upper
Mean	21.3975	1.95332	17.8920	25.5898

Log-Likelihood = -487.593
Goodness-of-Fit
Anderson-Darling (adjusted) = 18.202
Characteristics of Distribution

Table 8 Characteristics of distribution for maximum temperature (^{0}C)

	Estimate	Standard Error	95.0% Normal CI Lower	Upper
Mean(MTTF)	21.3975	1.95332	17.8920	25.5898
Standard Deviation	21.3975	1.95332	17.8920	25.5898
Median	14.8316	1.35394	12.4018	17.7375
First Quartile(Q1)	6.15568	0.5619	5.14721	7.36172
Third Quartile(Q3)	29.6632	2.70787	24.8036	35.4750
Interquartile Range(IQR)	23.5076	2.14594	19.6564	28.1133

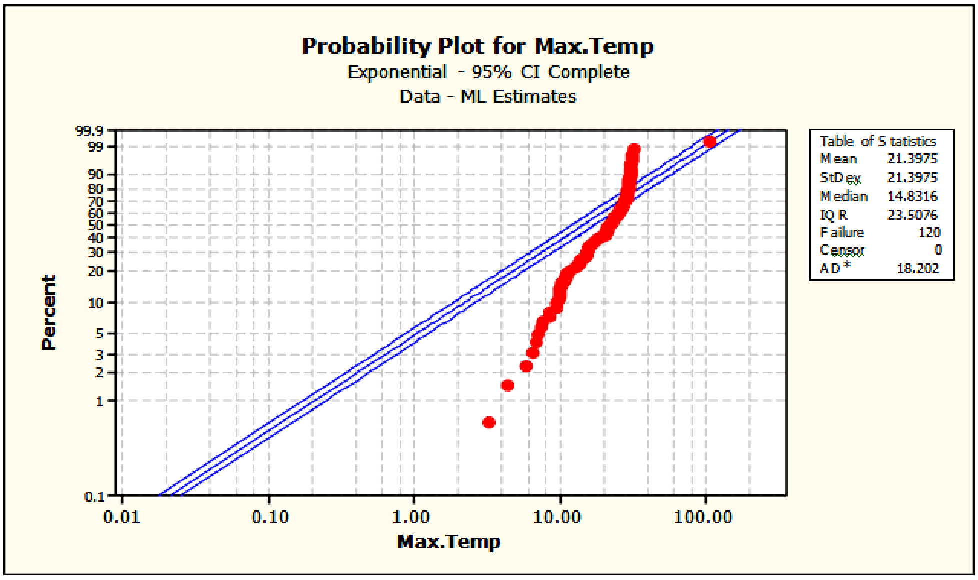

Fig 3 Probability plot for maximum temperature (^{0}C)

Distribution Analysis: R/H at 1730 hrs
Estimation Method: Maximum Likelihood
Distribution: Exponential

Table 9 Estimates for relative humidity (%)

Parameter	Estimate		95.0% Normal CI Lower	Upper
Mean	56.1833	5.12881	46.9790	67.1910

Log-Likelihood = -603.434
Goodness-of-Fit
Anderson-Darling (adjusted) = 39.542

Table 10 Characteristcs of distribution for relative humidity (%)

	Estimate	Standard Error	95.0% Normal CI Lower	Upper
Mean(MTTF)	56.1833	5.12881	46.9790	67.1910
Standard Deviation	56.1833	5.12881	46.9790	67.1910
Median	38.9433	3.55502	32.5634	46.5733
First Quartile(Q1)	16.1629	1.47547	13.5150	19.3297
Third Quartile(Q3)	77.8866	7.11004	65.1267	93.1465
Interquartile Range(IQR)	61.7237	5.63458	51.6117	73.8169

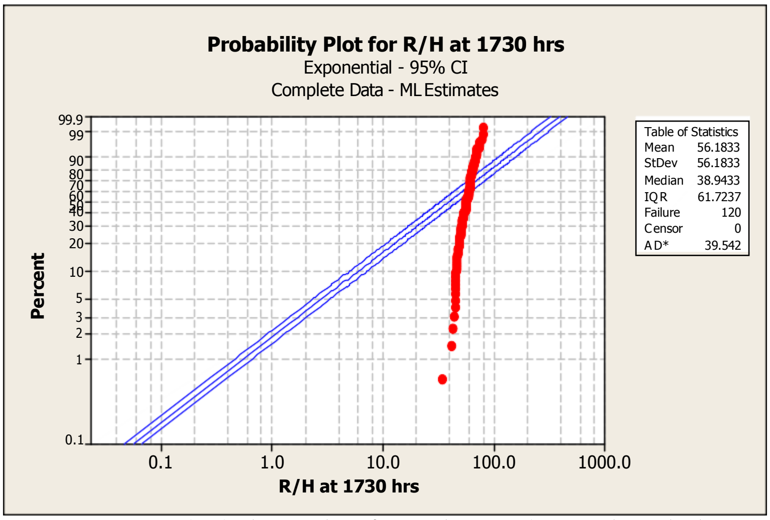

Fig 4 Probability plot for relative humidity (%)

The analysis of the data giving summary statistics(i.e mean, median, standard deviation, first quartile, third quartile, inter quartile range) and parameter estimation (i.e shape and scale) along with the graph of the probability plot shows that Weibull distribution fits well to the metrological data (i.e maximum temperature and relative humidity (%)).The estimation of the Statistical parameter particularly Anderson Darling Chi-Square is minimum through Weibull distribution for maximum temperature and relative humidity and the model is considered to be a good fit.

REFERENCES

BilaL, A., Raja, T.A. and Kumar, Parimal. 2005. On the Generalization of Weibull Distribution. SKUAST. *J.Res.*, **7**:201-207

Digest of Statistics.2006-2007. *Digest of Statistics. Directorate of Economics and Statistics*, Planning and Development Department, Government of Jammu and Kashmir.

Gilchrist, Warren 1984. *Statistical Modeling.* John Wiley and Sons. New York. Houghton Mifflin Company, Boston.

Hogg, R.V. and Craig, A.T. 1970. *Introduction to Mathematical Statistics*, Fourth edition, Macmillan, New York.

Johnson, N.L. and Kotz, S. 1970. *Continuous Univariate Distributions*, Vol 1 and Vol2 Lawless, J.F.

1982. *Statistical Models and Methods for Lifetime Data.* New York: John Wiley

Mudholkar, G.S., Srivastava, D.K. 1993. Exponentiated Weibull Family for analyzing bathtub failure-rate data. IEE Trans. *Reliability,* 42(2): 299-302.

Mudholkar, G.S., Srivastava, D.K., Freimer, M. 1995. *The Exponentiated Weibull Family: a reanalysis of the bus- motor-failure data. Tech.*, 37(4): 436-445.

Rao, C.R. 1973. Linear Statistical Inference and its Applications, Wiley Eastern Pvt. Ltd; New Delhi-India.