*Indian Journal of Agricultural Sciences* **81** (6): 533–8, June 2011

# Application of generalized lambda distribution for unimodal data

PRAJNESHU[1] and HIMADRI GHOSH[2]

*Indian Agricultural Statistics Research Institute, Library Avenue, Pusa, New Delhi 110012*

Received: 24 August 2009; Revised accepted: 4 April 2011

ABSTRACT

Pearsonian system of curves has been widely used to obtain the probability distribution of unimodal data sets. However, its main limitation is that each family requires different functional forms, which may be troublesome to implement especially on the boundaries between families. To this end, the purpose of this paper is to bring to the notice of Agricultural Scientists the existence of a very versatile family of generalized lambda distributions (GLD). A brief description is also provided of the recently developed GLDEX package, Ver. 1.0.3 in R, Ver. 2.8.1 software to fit GLD to data. An attractive feature of this software package is that it is freely downloadable. As an illustration, the probability density function of monthly rainfall data for Assam and Meghalaya meteorological subdivision is obtained.

**Key words**: Discretized method, Generalized lambda distributions, GLDEX package, Maximum likelihood method, Probability density function, Rainfall data, Starship method

Density estimation can give valuable indication of such features as skewness and modality in the data (Silverman 1986). The oldest and most widely used density estimator is the histogram. Given an origin $x_0$ and a bin-width h, the bins of the histogram are the intervals $[x_0 + mh, x_0 + (m + 1)h)$, where m is an integer. The intervals have been chosen closed on the left and open on the right for definiteness. The histogram is then defined as

$$\hat{f}(x) = (\text{Number of } X_i \text{ in same bin as } x) / (nh)$$

It may be noted that choice of the origin as well as bin-width is required for constructing the histogram. However, one limitation of this approach is that it is graphical and not analytical in nature. Pearsonian system of curves (Elderton and Johnson 1969) has been widely used to approximate unimodal data sets. However, its main limitation is that each family requires the solution of a different set of equations, which may be troublesome to implement especially on the boundaries between families. The Johnson distribution (Karian and Dudewicz 2000) is a four-parameter distribution based on the transformation of a standard normal variate. The Johnson distribution consists of three families and has been quite successful in practice. Its principle drawback is that it is not very easy to estimate the four parameters from the data.

[1]Principal Scientist and Head (e mail:prajneshu@yahoo.co.in), [2]Senior Scientist (e mail:hghosh@gmail.com), Division of Biometrics and Statistical Modelling

## MATERIALS AND METHODS

The generalized lambda distributions (GLD) (Karian and Dudewicz 2000) is a simple and flexible family of distributions that can assume a wide range of shapes. More importantly, GLD uses one general formula, unlike Pearsonian system of curves in which different Types (depending on the value of kappa) have different functional forms. The GLD's simplicity and versatility in fitting a broad range of curve shapes using only the first four moments, makes it an ideal choice. Flexibility of the GLD in assuming a wide variety of shapes has seen it being used extensively to fit and model a wide range of differing phenomena to continuous probability distributions, from applications in meteorology and modelling economic data, to Monte Carlo simulation studies.

A continuous probability distribution is defined in either of three ways, viz. by specifying its Probability density function, or through Distribution function, or through Percentile function (or Inverse distribution function). GLD is generally defined in terms of Percentile function (or Inverse distribution function). As an example, the Distribution function (D.f.) of N $(\mu,\sigma^2)$ is

$$F(x)=P(X \le x)=P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right)=\Phi\left(\frac{x-\mu}{\sigma}\right) \quad \ldots(1)$$

and the Percentile function (or Inverse distribution function) is

$$Q(y) = [x \text{ such that } F(x) = y] = \mu + \sigma\,\Phi^{-1}(y)$$

The four-parameter GLD (Karian and Dudewicz 2000) is given by

$$Q(y) = Q(y;\lambda_1, \lambda_2, \lambda_3, \lambda_4) = \lambda_1 + \left\{ y^{\lambda_3} - (1-y)^{\lambda_4} \right\} / \lambda_2, \; 0 \le y \le 1 \quad \ldots(2)$$

The parameters $\lambda_1$ and $\lambda_2$ are respectively location and scale parameters while $\lambda_3$ and $\lambda_4$ determine skewness and kurtosis. However, one limitation of Eq (2) is that the functional form is valid only for some particular values of $\lambda_i$, i=1,2,3,4. Accordingly, Freimer *et al*. (1988) proposed another parameterization (referred to as FMKL), which is valid for all the values of $\lambda_i$, i=1,2,3,4. It is specified by the percentile function Q(y):

$$Q(y) = \lambda_1 + \frac{1}{\lambda_2}\left[ \frac{y^{\lambda_3} - 1}{\lambda_3} \;\; \frac{(1-y)^{\lambda_4} - 1}{\lambda_4} \right], \; 0 \le y \le 1 \quad \ldots(3)$$

Noticing that probability density function (pdf) f(x) satisfies

$$f(x) = [dQ(y)/dy]^{-1},$$

Eq (3) gives at x = Q(y):

$$f(x) = \lambda_2 / \{ y\lambda_3 - 1 + (1-y)\lambda_4 - 1 \} \quad \ldots(4)$$

GLD family includes a large number of well-known distributions, like normal, uniform, exponential, gamma, Weibull, lognormal, beta, chi-square, inverse Gaussian, logistic, Pareto, and F. Details are given in Karian and Dudewicz (2000). It may be pointed out that GLD family is much richer than Pearsonian family, as, for example, logistic distribution belongs only to former.

*Shapes of GLD*

The variety of shapes offered by GLD are as follows:

Class I. $\lambda_3<1$, $\lambda_4<1$: Unimodal densities with continuous tails. This class can be subdivided with respect to the finite or infinite slopes of the densities at the end points.

Class II. $\lambda_3>1$, $\lambda_4<1$: Monotone pdfs similar to those of the exponential or $\chi^2$-distributions. The left tail is truncated.

Class III. $1<\lambda_3<2, 1<\lambda_4<2$: U-shaped densities with both tails truncated.

Class IV. $\lambda_3 > 2$, $1<\lambda_4<2$: Rarely occurring S-shaped pdfs with one mode and one antimode. Both tails are truncated.

Class V. $\lambda_3 > 2, \lambda_4 > 2$: Unimodal pdfs with both tails truncated.

*Estimation of parameters of GLD*

*Method of moments*: The most popular method for estimating the parameters of GLD is to match the first four moments of the empirical data to those of the GLD. The popularity of this method is partly due to the availability of extensive tables that provide parameter values for given values of skewness and kurtosis (Karian and Dudewicz 2000). However, care must still be taken as different parameter values can return the same moments and hence the ensuing GLD may fail to properly represent the actual distribution of the data.

If the mean, variance, skewness, and kurtosis of the data are respectively $\mu^*$, $\sigma^{*2}$, $\alpha^*_3$, and $\alpha^*_4$, the parameter estimates of $\lambda_3$ and $\lambda_4$ of the GLD can be obtained by solving (Lakhany and Mausser 2000):

$$\alpha_3 = \alpha^*_3 \text{ and } \alpha_4 = \alpha^*_4 \quad \ldots(5)$$

Once $\lambda_3$ and $\lambda_4$ are estimated, the parameters $\lambda_2$ and $\lambda_1$ can respectively be estimated as:

$$\lambda_2 = (v_2 - v_1^2)^{1/2} / \sigma_* \quad \ldots(6)$$

$$\lambda_1 = \mu^* + \frac{1}{\lambda_2}\left( \frac{1}{\lambda_3 + 1} + \frac{1}{\lambda_4 + 1} \right) \quad \ldots(7)$$

where $v_i$ is the $i^{th}$ raw moment. Unfortunately, exact solutions to (5) do not exist and there is a need to use numerical methods to obtain approximate solutions. This involves using optimization techniques to find $\lambda_3$ and $\lambda_4$ such that

$$\left(\alpha_3^* - \alpha_3\right)^2 + \left(\alpha_4^* - \alpha_4\right)^2 < \varepsilon \quad \ldots(8)$$

where ε is a positive number representing the desired accuracy.

*Method of maximum likelihood:* Su (2007) suggested a two-step procedure using the method of moments to find initial values and then maximizing the numerical log likelihood to fit the appropriate GLD to data. The algorithms can be summarized as follows:

1. Specify a range of initial values for $\lambda_3$, $\lambda_4$, and the number of initial values to be selected. Here, the $\lambda_3$, $\lambda_4$ are set by default to the range from -0.25 to 1.50 for the method of moments. These default values appear to work well in most situations. However, it is possible to change these, if desired. Scrambling methods are applied so that the numbers generated fills uniformly onto the $\lambda_3$, $\lambda_4$ two-dimensional space. Alternatively, other common quasi-random number generators, such as Halton and Sobol sequences (available in R) can also be used. By default, 10000 of such initial values are chosen and used in Step 2.
2. Evaluate $\lambda_1$, $\lambda_2$ for each of the initial values $\lambda_3$, $\lambda_4$. Remove all the sets of initial values that do not:
   (a) Result in a legal parameterization of GLD or
   (b) Span the entire region of the data set.
   Among the sets of initial points not excluded by Step 2, find a set of initial values that produce the lowest value in expression (8). This set of initial values will then be used in the optimization process.
3. Calculate quantiles $y_i$ from the initial values of GLD. This can be done by solving expression (5) numerically.
4. Once $y_i$ is obtained, substitute it into the following numerical log likelihood (9), obtained by using the chain rule to differentiate expression (5) to obtain $f(x_i)$

and applying the logarithm on the joint distribution of $f(x_i)$, assuming their independence:

$$ML = \sum_{i=1}^{n} \log \left[ \frac{\lambda_2}{y_i^{\lambda_3 - 1} + (1-y_i)^{\lambda_4 - 1}} \right], \qquad \ldots(9)$$

5. The optimal result can be obtained via the Nelder-Mead Simplex algorithm or any another suitable numerical optimization algorithm. It is always desirable to find another set of initial values in the optimization process to check as to whether or not the result obtained is a reasonable solution. The final fitting result can be examined by plotting the histogram with the fitted line and quantile plot as well as testing the goodness of fit using the resample Kolmogorov-Smirnov (KS) tests.

*Discretized Method*: Su (2005) proposed a discretized approach, which is non-parametric, to flexibly fit GLD to data. It optimizes the bin-width of data histogram to find a suitable GLD. In addition to the default optimization, this approach provides additional flexibility akin to the concepts of 'Loess' and 'Kernel smoothing', which allow the users to determine the amount of details they would like to smooth over the data. Description of the algorithm is as follows:

1. Specify a range of initial values for $\lambda_3$, $\lambda_4$, and the number of initial values to be selected. Here, the $\lambda_3$, $\lambda_4$ are set by default to the range from –0.25 to 1.50 for the method of moments. These default values appear to work well in most situations. However, it is possible to change these, if desired. Scrambling methods are applied so that the numbers generated fills uniformly onto the $\lambda_3$, $\lambda_4$ two-dimensional space. To increase the speed, it is possible to set the initial values where $\lambda_3 = \lambda_4$. This appears to work well in many situations. By default, 100 of such initial values are chosen and used in Step 2.
2. Evaluate $\lambda_1$, $\lambda_2$ for each of the initial values $\lambda_3$, $\lambda_4$. Remove all the sets of initial values that do not:
   (a) Result in a legal parameterization of GLD or
   (b) Span the entire region of the data set.
   From these sets of initial points, find the values of $\lambda_3$, $\lambda_4$ that matches closely with the data. This is to generate a set of initial values that produce the lowest value in Expression (8). This set of initial values will then be used in the optimization process.
3. Sort the data in ascending order, and divide the data set into evenly spaced classes with bin edges that span the data set. Calculate the proportion of the sample out of the total sample in each class. Let the proportion of data in each class be denoted by $d_i$, i=1,2,…,n. And the proportion of data from the theoretical GLD be the vector $t_i$, i=1,2,…,n. The quantity that is to be minimized is:

$$\sum_{i=1}^{n} d_i \left(d_i - t_i\right)^2 \qquad \ldots(10)$$

The expression (10) involves weighted squared deviations so that the data with higher proportions are given priority in the minimization scheme. The philosophy is to choose the number of classes, n, that best represents the first two moments of the data.

4. The optimal result can be obtained via the Nelder-Mead Simplex algorithm or any another suitable numerical optimization algorithm. It is advisable to re-use the initial values in the optimization process to ensure that the result obtained is a global minimum rather than a local minimum. Steps 1 to 3 may be repeated, if necessary, where the number of classes and the range of initial values can be adjusted until the results are deemed adequate. The final fitting result can be examined by plotting the result on the histogram with the fitted line as well as testing the goodness of fit using the Kolmogorov-Smirnov (KS) test.

*Starship Method:* King and Mac Gillivray (1999) developed a computer intensive estimation method for fitting GLD, called "Starship method". This method involves three steps:

1. For a set of data and a range of $\lambda_1$, $\lambda_2$, $\lambda_3$ and $\lambda_4$ values, apply the reverse transformation, i.e. a data value X is transformed to F(X). (note that as F does not exist in closed form for the GLD, numerical methods are needed)
2. Calculate the value of a suitable goodness of fit measure for the closeness of the resulting values to the Rectangular R[0,1] distribution
3. Choose the ($\lambda_1$, $\lambda_2$, $\lambda_3$ , $\lambda_4$) values that minimize the chosen goodness of fit measure to the Rectangular distribution, as the fitted values

After a good deal of computations, it was found that the two-stage process with second steps half the first, i e 10 × 10 × 10 × 10 grid gives the most efficient balance of computational speed and accuracy of results. An advantage of this method is that it can easily incorporate the alternative measures of goodness of fit to the Rectangular distribution, like the Kolmogorov distance, and the Anderson-Darling distance, when exploring fits. The following method is used to select the region of parameter space that is searched:

(a) The location parameter, $\lambda_1$, ranges from the first to the third quartile of the data.
(b) The scale parameter, $\lambda_2$, is chosen to match with the ranges of the shape parameters so that reasonable values for, say $F^{-1}(0.01)$ and $F^{-1}(0.99)$ are included in the searched area.
(c) The shape parameter ranges are chosen by considering the apparent shape of the data or distribution and referring to the section that describes the type of the distributions for different $\lambda_3$ and $\lambda_4$ values in Freimer *et al.* (1988).

*GLDEX software package*

This is a very versatile software package and is designed to fit GLD to data. The GLDEX package, Ver. 1.0.3 in R, Ver. 2.8.1 software was released in December, 2008. It is a freeware and is available from the Comprehensive R Archive Network at http://CRAN.R-project.org/. The following command loads the package and allows browsing of a summary of the important functions of this software package:

*R> library( "GLDEX" )*

*R> ?GLDEX*

For unimodal data, GLDEX provides the Maximum likelihood estimation (Su 2007) as well as the Discretized nonparametric method (Su 2005), which acts as a smoothing device similar to the concept of kernel density estimation or loess smoothing. The GLDEX package also includes the Starship method (King and MacGillivray 1999). Further, for assessing goodness of fit, several methods, like the histograms, quantile plots and Resample Kolmogorov-Smirnov (KS) test are included in the package.

*Fitting GLD to data*

To fit GLD to a data set, it is necessary to find: (a) suitable initial values, and (b) optimize the values through an optimization scheme. The initial values and the optimization scheme required for each method are discussed below:

*Finding initial values:* The first step is to generate a set of feasible initial values. For the FMKL parameterization of GLD, the initial values of $\lambda_3$ and $\lambda_4$ comprise low discrepancy quasi-random numbers ranging from –0.25 to 1.5. These values are chosen as they appear to work well for a wide range of situations and can be modified, if necessary. Once generated, these values can be used to estimate $\lambda_1$ and $\lambda_2$ using the method of moments. From this set of initial values, GLDEX will attempt to find the best set of initial values for subsequent optimization process. The goal is to find the GLD that matches most closely with the third and fourth moments of the data, using the minimum squared criterion.

*Terminology:* The discretized approach is known as the revised method of moment FMKL (RMFMKL) method. The maximum likelihood approach adds a suffix .ML to the name, so the method is labeled as RMFMKL .ML. These abbreviations are used frequently in the GLDEX package and form a part of the graphical outputs to allow distinction between different fitting methods.

*Discretized approach:* Here, the data is sorted in ascending order and divided into evenly spaced classes with bin edges that span the data set. Then the proportion of the sample in each class is calculated. As an example, Table 1 shows 4 classes, with the proportion of the data set belonging to each class shown in the 2nd row. For i=1,2,…,n classes, the proportion of data in each class is defined as $d_i$ and the proportion of data from the GLD is $t_i$.

Table 1 Proportion of data in each class

| Classes | 1.5-2.0 | 2.0-2.5 | 2.5-3.0 | 3.0-3.5 | Sum |
|---|---|---|---|---|---|
| Proportion of data | 0.1 | 0.6 | 0.2 | 0.1 | 1.0 |

The quantity to be minimized, indicated in expression (10), is the weighted squared deviation of theoretical proportions from empirical proportions. This weighting scheme forces data with higher proportions to be given priority in the minimizations scheme and this tends to accentuate the peak and suppress the tails of the empirical data. The weighting factor $d_i$ can be removed resulting in expression (11):

$$\sum_{i=1}^{n}\left(d_i - t_i\right)^2 \quad ...(11)$$

The number of classes, n, can be solely determined by the user, or determined by finding the number of classes that best matches the mean and variance of the actual data set in terms of minimum squared error.

*Maximum likelihood estimation:* Here, it is necessary to obtain quantiles $u_i$ for every observation $x_i$, i=1,2,…,n under a set of initial values. This requires solving Eq (3) numerically using, say, the Newton-Raphson method, available in the package. Once the $u_i$'s are obtained, these are substituted in Eq (9). The key here is to maximize the likelihood in Eq (9) and this can be done using Nelder-Simplex algorithm. To check the numerical optimization, it is always desirable to use a different set of initial values to see if similar results can be obtained in the optimization process.

*Assessing the quality of fit*

This can be done by using three methods available in GLDEX package, as discussed below:

*Graphical outputs:* The most obvious diagnostic check on the resulting distribution fit is to superimpose the resulting distribution fit onto the histogram. While simple and effective, it has shortcomings as it can be difficult to assess the adequacy of the distributional fit on the tails and it remains subjective matter as to what constitutes a good or bad fit. Different classes or number of bins in the histogram can also give different distributional shape of the data set. It may not be easy to determine whether the resulting fit is adequate if it appears to capture the shape of the data very well under a histogram with 10 bins but not so well with 50. For this reason, quantile plots are also provided so that the user can see more objectively as to in which part of the data, the GLD distribution appears to give an adequate fit.

*Comparing the mean, variance, skewness and kurtosis of the fitted distribution with the empirical data:* This method provides a more objective way of choosing between alternative distributional fits. The derivation of the four moments of the FMKL parameterization of GLD involves the use of beta function. It is imperative to appreciate, however, a GLD that has very similar mean, variance, skewness and kurtosis to the actual data may still be a bad fit (Karian and Dudewicz 2000, Lakhany and Mausser 2000).

In some cases, it may be desirable to choose a good distributional fit with the closest mean, variance, skewness and kurtosis to the data set so that the fitted distribution can be used for simulation studies to model the population of interest.

*Resample Kolmogorov-Smirnov (KS) test:* This test assesses the similarity between fitted distribution and actual data by sampling a proportion (for example 90%) of the data and fitted distribution and calculating the KS test p-value. This process is then repeated many times, and the number of times the p-value is not significant is recorded and reported. For example, if 950 times out of 1000 times, the p-value does not reject the null hypothesis, it is possible to state that it is quite likely that the resulting fit is quite adequate for the given data set.

## RESULTS AND DISCUSSION

The monthly rainfall data during 1871-2006 in respect of all the meteorological subdivisions of the country are available at the website *www. tropmet.res.in* of the Indian Institute of Tropical Meteorology, Pune. As an illustration, the data pertaining to Assam and Meghalaya meteorological subdivision for the months of summer monsoon rainfall, i.e. for the months of June, July, August, and September every year are considered for data analysis. Thus, the total number of data points is 544 and file name for the data is datau[,1]. The first step is to assess as to whether this data set has any mode, or is unimodal, or is multimodal. To this end, Density estimation can give valuable indication of such features as skewness and modality in the data (Silverman 1986). It may be noted that choice of the origin as well as bin-width is required for constructing the histogram. The following command available in GLDEX, Ver. 1.0.3 software package was used to construct histograms for the data for various choices of the origin and bin-widths:

h1<- hist.su(datau[,1])

The histograms give an indication that the underlying distribution is unimodal. However, one limitation of this approach is that it is graphical and not analytical in nature

Subsequently, attempts were made to apply GLD to the data by three methods, viz Discretized method, Maximum likelihood method, and Starship method available in GLDEX, Ver. 1.0.3 software package by employing respectively the following commands:

Table 2 Fitting of GLD to Assam and Meghalaya monthly rainfall data

| Parameter | Methods | | |
|---|---|---|---|
| | Discretized | Maximum likelihood | Starship |
| $\lambda_1$ (location) | 3650.43 | 3654.96 | 3658.50 |
| $\lambda_2$ (scale) | 0.0016 | 0.0013 | 0.0013 |
| $\lambda_3$ (shape) | 0.2688 | 0.2803 | 0.2671 |
| $\lambda_4$ (shape) | 0.1513 | 0.0571 | 0.0482 |

Table 3 Goodness of fit of GLD

| Parameter | Data | Methods | | |
|---|---|---|---|---|
| | | Discretized | Maximum likelihood | Starship |
| *Comparison of various statistics* | | | | |
| Mean | 3783.44 | 3701.07 | 3781.33 | 3783.29 |
| Variance | 1207073 | 696140 | 1205808 | 1213335 |
| Skewness | 0. 52 | 0.23 | 0.52 | 0.53 |
| Kurtosis | 3.33 | 2.78 | 3.39 | 3.47 |
| *Resampled K-S test* | | | | |
| *p-value (out of 1000)* | – | 328 | 967 | 961 |

fit1<- fun.data.fit.hs (datau[,1])
fit2<-fun.data.fit.ml(datau[,1])

and the results obtained are reported in Table 2.

Finally, the goodness of fit is examined by several methods. Firstly, the mean, variance, skewness and kurtosis of the fitted GLD are compared with the corresponding values for the data by employing the following commands:

fun.theo.mv.gld(fit1[1,2],fit1[2,2],fit1[3,2],fit1[4,2],param="fmkl")
fit3<-fun.comp.moments.ml(fit2,datau[,1],name="ML")

and the results are reported in Table 3. Subsequently, the simulated Kolmogorov-Smirnov test statistics for the fitted GLD are computed by employing the following commands:

fun.diag.ks.g(fit1[,2],datau[,1],param="fmkl")
fun.diag2(obj.fit1u.ml,datau[,1],1000)

and the results are reported in Table 3. A perusal indicates that the maximum likelihood method has performed the best. To get a visual insight, the graph of quantile (qq-plot) for the fitted GLD through maximum likelihood method is obtained by employing the following command:

qqplot.gld(datau[,1],obj.fit1u.ml[,2],param="FMKL", name="ML Fit")

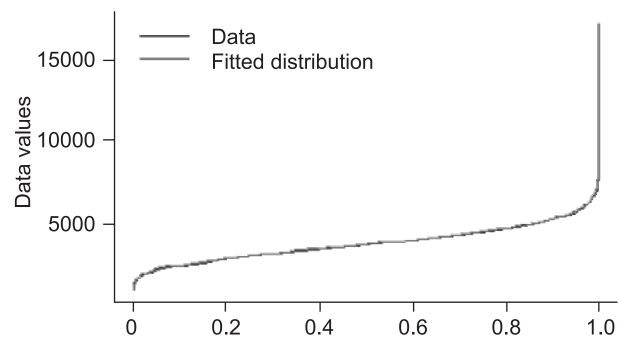


Fig 1 Quantile plot for the data and fitted GLD using maximum likelihood method

and the same is exhibited in Fig 1. A perusal indicates that, for given data, the fit of GLD by maximum likelihood method is very good.

Finally, employing the command:

plot1<-fun.plot.fit(fit2,datau[,1],nclass=50, param=c

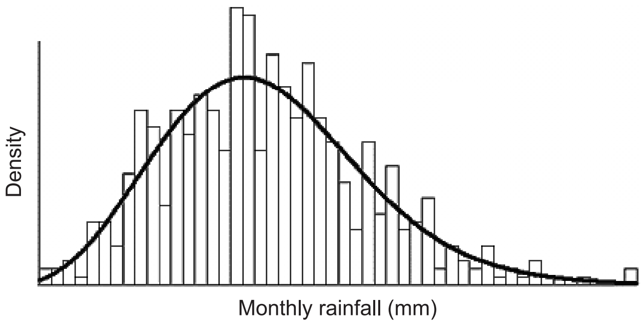


Fig 2 Graph of fitted GLD by maximum likelihood method along with histogram of data

("fmkl"),xlab="datau[,1]")
the graph of fitted model by maximum likelihood method along with histogram of data is exhibited in Fig 2. Evidently, the two are seen to be quite close to each other.

The procedure for estimation of probability distribution of unimodal rainfall data using Generalized lambda family is thoroughly discussed. As future work, there is a need to develop the methodology to obtain "thresholds" of rainfall by using the knowledge of the probability distribution obtained. This type of effort would go a long way in putting Rainfall-based crop insurance on a sound statistical footing.

## REFERENCES

Elderton W P and Johnson N L. 1969. *Systems of Frequency Curves.* Cambridge University Press U.K.

Freimer M, Mudholkar G, Kollia G S and Lin C T. 1988. A study of the generalized Tukey lambda family. *Communications in Statistics, Theory and Methods* **17**: 3547-67.

Karian Z and Dudewicz E. 2000. *Fitting Statistical Distributions: The Generalized Lambda Distribution and Generalized Bootstrap Methods*. CRC Press Boca Raton.

King R A R and Mac Gillivray H L. 1999. A starship estimation method for the generalized λ distributions, *Australia and New Zealand Journal of Statistics* **41**: 353-74.

Lakhany A and Mausser H. 2000. Estimating the parameters of the generalized lambda distribution. *Algo Research Quarterly* **3**: 47-58.

Silverman B W. 1986. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall U. K.

Su S. 2005. A discretized approach to flexibly fit generalized lambda distributions to data. *Journal of Modern Applied Statistical Methods* **2**: 408-24.

Su S. 2007. Numerical maximum log likelihood estimation for generalized lambda distributions. *Computational Statistics & Data Analysis* **51**: 3983-98.