Pak. J. Statist.
2007 Vol. 23(2) pp 157-169

FITTING PEARSON III AND LOG-PEARSON III DISTRIBUTIONS ON FLOODS DATA OF INDUS RIVER AT KOTRI BARRAGE

Abdul Ghani Memon
Department of Statistics
Shah Abdul Latif University
Khairpur, Mir's

ABSTRACT

This study has been undertaken for the purpose of prediction of floods of Indus River at Kotri barrage, by fitting Pearson III (P-III) and Log – Pearson III (LP-III) distribution on the data of 99 years. The Quantiles and their Standard Errors (S.E's) are estimated, for each distribution, using Method of Moments (MOM), Maximum Likelihood Method (MLM) and Probability Weighted Moments (PWM). The results are tested using χ^2 – test, S-K test, Probability Plot Correlation Co-efficient (PPCC), Root Mean Square Errors (RMSE), Co-efficient of Skewness (C_S), Co-efficient of Kurtosis (C_K) and Sample L-moments.

It has been observed that P-III is better than LP-III and MLM is more efficient than MOM. A flood of 0.91 millions cusecs is expected to pass through Kotri barrage by fitting P-III, using MLM, during next 100 years. The expected flood being higher than the designed capacity by 36 thousand cusecs suggests that there is need to take appropriate measures to save the structure of barrage from any damage.

KEY WORDS

Pearson III, Log – Pearson III, Method of Moment, Maximum Likelihood Method, Indus, Kotri

1. INTRODUCTION

The estimates of floods of specified frequency beyond the recorded range play an important role in the design of hydraulic structures. These estimates are made through the use of a particular probability distribution function fitted to a historical data.

Forecast of floods can generally reduce the dangers caused by floods if proper probability distribution and an efficient method of estimation is used for forecasting. Identification of a probability distribution of flood data is connected to the necessity of characterizing in a synthetic form the data collected and the provision of tools for making statistical analysis.

Once a distribution is selected for use, the analysis becomes strictly objective and the extrapolation is automatically made from a mathematical determination of the sample Statistic, i.e. mean, S.D and C_s. Different statisticians and hydrologist have proposed and used various distribution and estimation methods, for prediction of flood of rivers.

The first scientific study of the data of Indus river, at Kotri, from 1901 to 1943 and from 1901 to 1976, was carried out by Nixon mentioned in Bund Manual (1977). Pearson Type-I skew curve (1977) was fitted and it was noted that a flood of 0.7 million cusecs and 0.6 million cusecs or more may be expected with a frequency of less than 5 in 1000 years and less than 2 in a century (43 years data) and that lowest flood of 0.19 million cusecs and highest flood of 1.05 million cusecs have no chance of occurrence (76 years data).

Design capacity of the barrage, constructed in 1955, was 0.815 million cusecs. In 1956, 0.981 million cusecs had actually passed and more than 0.7 million cusecs had actually passed at Kotri six times. Prediction by Gumbel , Log- normal, Foster-III, Pearson Type- I and actual occurrence of floods suggest remodeling of the barrage. Nixon, Memon,. A.G had fitted various probability distributions on the data of Indus river at Kotri for different sample sizes. Their results are summarized below:

Designed Capacity 8.75 million cusecs (1955) Kotri

Sample Size	Distribution/ Curve	Flood estimates for 100 years (Million Cusecs)	Investigators
43	Pearson Type-I	0.75	Nixon (1977)
76	Pearson Type –II	0.675	Nixon (1977)
99	Gumbel (MLM)	0.975	Memon & Shaikh (2005)
99	Log Normal(MOM)	0.95	Memon (2007)

Matales and Wallis (1973) comparing the quantile estimates by P-III, using MOM and MLM, observed that MLM estimates were less biased with smaller Standard Error (S.E). The differences between MOM and MLM are larger for smaller samples.

Bucket and Oliver III compared MOM and MLM for P-III and recommended MLM. PWM was applied to P-III by Singh, V.P and Singh, K (1988) and they found it as efficient as MOM. Shaligran and Lele (1978) had analyzed flood flows from 16 streams in India using P-III.

In this study, as in (2005), P-III and LP-III have been fitted on the data of flood peaks at Kotri, from 1901 to 1999. LP-III has been recommended by water Reserves Council of USA and used by Askhar and Bobee LP-II, Arrora and Singh (1989).

The data of Peak floods for Kotri barrage was collected from the office of the Chief Engineer, Kotri barrage at Hyderabad. Histogram is presented in Fig.1 Historigram and frequency curve are presented in Fig.2 95% Confidence Intervals for Quantiles estimates by P-III and LP-III distributions using MLM along with the observed and estimated floods are given in Fig. 3 and Fig. 4 respectively. Fig.5 is the plot of P-III and LP-III by MLM. Mean, Median, Mode S.D, C_s, C_k , S.E's, Sample L-moments, Ist Quartile (Q_1), 3 rd Quartile (Q_3), etc, are presented in Table I. SPSS, Excel, Minitab and Data plot software are used for Statistical analysis .

2. METHODOLOGY

The Probability distributions are fitted and the results are tested and compared step-wise as follows:

i) Parameters of each distribution (by three methods of estimation) along with quantile estimates and their S.E's, for different return periods, are calculated.
ii) Various goodness of fit tests, $C_{s,}$ $C_{k,}$ etc, are used.
iii) Ratios, Differences and Confidence Intervals are also displayed by tables and figures.

Theoretical details are given in (8) and application for the fitting of P-III distribution by PWM to annual peak flood discharge data at Kotri is demonstrated in appendix.

3. RESULTS AND DISCUSSION

i) Descriptive Statistics (Table 1)

Fig. 1 shows that high floods of 0.981, 0.791, 0.722, 0.825, 0.799, 0.804, 0.764 (million cusecs) have occurred in the years 1956, 1976, 1978, 1994, 1995, 1995, 1958, respectively.

Fig 2 shows that Super floods (higher than 0.7 million cusecs) have occurred 11 times and high floods (more than 0.6 million cusecs) have occurred 17 times.

C_s is 0.83 and C.V is 0.36.
Mean > Median > Mode. Thus the data is positively skewed.
The value of each Statistic increases as the sample size increases.

ii) Parameter and Quantile Estimates (Table 2 and 3)

Parameters estimated by P-III and LP-III, using three method of estimation, are presented in Table 2 and Table 3.
The Quantile estimates are compared in Table 4 A, and discussed as under:

(a) P-III

$Q_{(MOM)}/Q_{(MLM)}$ varies from 1.002 to 1.008

$Q_{(PWM)}/Q_{(MLM)}$ varies from 1.01. to 1.04

The differences between estimates by MOM and MLM are very small.
MLM gives smaller estimates of Quantiles than MOM.

(b) LP-III

$Q_{(MLM)}/Q_{(MOM)}$ varies from 1.02 to 1.10

$Q_{(PWM)}/Q_{(MOM)}$ varies from 1.007 to 1.03

The estimates by PWM are nearly equal to those by PWM.
MOM gives the smallest quantile estimates.
Thus the estimates of quantiles by L P-III are different than by P-III.

(c) P-III Versus LP-III

The ratio varies between 1.002 and 1.005, 0.90 and 0.98, 1.002 and 1.01, respectively, by MOM, MLM and PWM. Thus P-III by MLM gives smallest quantiles estimates.

Note. 0.91 and 1.00 million cusecs flood is expected to pass at Kotri during next 100 years, respectively, by P-III and LP-III, using MLM.

iii) Models of P-III and LP-III Distributions

Table-5
Models of P-III and LP-III Distributions

Distribution	Barrage	Method	Models
P-III	Kotri	MOM	$\hat{x}_T = (66.2 \times 5.81) + (76.5) + K_T \sqrt{(66.2)^2 \times 5.81}$
		MLM	$\hat{x}_T = (62.2 \times 6.49) + (47.5) + K_T \sqrt{(6.22)^2 \times 6.49}$
		PWM	$\hat{x}_T = (80.5 \times 4.08) + (123.5) + K_T \sqrt{(80.5)^2 \times 4.08}$
LP-III	Kotri	MOM	$\hat{x}_T = \exp[(0.039 \times 82.95) + (2.81) + K_T \sqrt{(0.039)^2 \times 82.95}\ ]$
		MLM	$\hat{x}_T = \exp[(0.0191 \times 382.5) + (-0.55) + K_T \sqrt{(0.0191)^2 \times 382.5}\ ]$
		PWM	$\hat{x}_T = \exp\ [(0.0194 \times 338.25) + (-0.44) + K_T \sqrt{(0.0194)^2 \times 338.25}\ ]$

iv) S.E's of Quantile estimates (Table 3 and 4 B)

a) P-III: $SE_{(MOM)} / S.E_{(MLM)}$ varies between 1.02 to 1.06, i.e., MLM is efficient than MOM.

b) LP-III: $SE_{(MOM)} / SE_{(MLM)}$ varies between 1.01 to 1.04 , i.e., MLM is more efficient than MOM (as in case of P (3)).

c) P-III .Versus LP-III

MOM: The ratio varies between 0.97 and 0.999.
Thus P-III is better than LP-III
MLM: The ratio varies between 0.94 and 0.996.

Thus P-III is better than LP-III, as by MOM.
P-III gives smaller S.E's of quantiles than LP-III, both by MOM and MLM.
The ratios are smaller by MLM than by MOM.

Hence, P-III is a better fit than LP-III , specially by MLM.

Fig 3 and Fig 4 are the plots of P-III and LP-III distributions curves by MLM with confidence bands of 95%. Both curves show that computed quantiles (X_T) of both the distributions by MLM method give a better fit by forming a straight line through the mean point (T = 2.33, $\overline{X}$ = 0.451million cusecs) and within the confidence bands of X_1 (95%) and X_2 (95%).
Fig 5 shows that both the distributions are acceptable but P-III is superior than LP-III as it passes exactly through the mean point.

v) Goodness of Fit Tests

Table-6
Tests of Goodness of Fit (Kotri) 1901-1999

χ^2 –test			S-K Test			PPCC		RMSE	
α=5%	χ^2_{cal}		α =5%	D_{cal}					
	P-III	LP-III		P-III	LP-III	P-III	LP-III	P-III	LP-III
5.990	4.85	4.24	0.136	0.128	0.112	0.9946	0.9928	0.0566	0.0691

Above Table shows that both the distributions are accepted on the basis of all tests. Moreover, P-III is better than LP-III on the basis of PPCC and RMSE.

vi) C_s, C_k and L-Moment Ratios

Table-7
Values of C_s, C_k, t_3 and t_4 (Kotri) 1901-1999

	Observed (using Data)	**Values Suggested for Pearson III**	**Appx. Values obtained for Log- Pearson III**
t_3	0.163	0.163	0.165
T_4	0.149	0.131	0.133
C_s	0.830	0.830	0.842
C_k	3.537	4.03	4.147

Above table indicates that both the distributions are acceptable on the basis of t_3, t_4, C_s and C_k. P-III is better than LP-III on the basis of t_4 and C_k.

4. CONCLUSION

i) The data is positively skewed.
Mean > Median > Mode (as for Guddu and Sukkur Barrages).
ii) Quantiles estimates are smallest for LP-III by MLM and for P-III by MLM.
P-III by PWM gives the largest quantile estimates.
iii) MLM is more efficient than MOM, specially for P-III.
iv) P-III and LP-III distributions are accepted on the basis of tests of significance and L- moment ratios. P-III is better than LP (3) on the basis of PPCC,RMSE, t_4 and C_k.
v) 0.91 millions cusecs flood is expected to pass at Kotri barrage, during next 100 years, by fitting P-III using MLM.

RECOMMENDATIONS

It is recommended to fit P-III distribution using MLM Parameter estimation method for prediction of occurrence of individual peak floods on the Indus river at Kotri barrage. The predicated flood by P-III distribution, using MLM method, is 0.91 million cusecs which is much higher than the designed capacity of Kotri barrage. Thus there is an urgent need to take the appropriate measures to save the structure of barrage from occurrence of any possible damage as the above predicted flow can occur at any time due to environmental changes.

ACKNOWLEDGEMENT

The author is thankful to the reviewer for useful comments.

REFERENCES

1. Arora, K. and Singh, U.P (1989). A Comparative Evaluation of Estimators of the Log Pearson (3) Distribution, *Journal of Hydrology*, Vol. 105, 19-37.
2. Ashkar, F., and Bobee, B. (1988). Confidence Intervals for Flood Events Under a Pearson (3) or Log- Pearson–(3) Distribution. *Water Res. Bulletin*, Vol. 24. No. 3, 639-650.
3. Buckett, J. and Oliver, F.R. (1977). Fitting the Pearson Type 3 Distribution in Practice. *Water Resources Research*, Vol. 13, No. 5, 851-852.
4. Bund Manual (1977). *Irrigation and Power Department*. Government of Sindh.
5. Matalas, N.C. and Wallis. J.R. (1973). Eureka. It fits a Pearson Type 3 Distribution, *Water Resources Research*, Vol. 9, No. 2, 281-289.
6. Memon, A.G. and Shaikh, N.M. (2005). Frequency Prediction Analysis of Flood Data of Indus River at Kotri Barrage by Gumbel and Generalized Extreme Value Distributions, Mehran University. *Research Journal of Engineering and Technology*, Vol. 24(1), 23-34.
7. Memon, A.G. (2007). A comparison of Normal and Log–Normal (2) Distributions Using Three Estimation Methods On Annual Flood Peaks Data of River Indus at Kotri Barrage, *Pak. J. Statist*, Vol. 23 (1), 37-46.
8. Ramachandra Rao, A (2000). *Flood Frequency Analysis*. CRC Press, London.
9. Shaligram, V.M. and Lele, V.S. (1978). Analysis of Hydrologic Data using Pearson Type III Distribution. *Nordic Hydrology*, Vol. 9, 31-42.
10. Sing, V.P. and Singh, K. (1988). Parameter Estimation for Log-Pearson Type III Distribution. *Journal of Hydraulic Engineering*, ASCE, 112-122.

Table-1
Basic statistics (Kotri Barrage) 1901-1999

N 99 (85)	Mean* 451.95(444.0)	Median* 420.0(429.0)	Mode* 356.70(360.0)
Min* 153.0 (155.0)	Max* 982.0 (982.0)	C_s 0.830 (0.82)	C_k 3.537 (3.50)
S.D 159.645 (161.6)	Q.D 104 (108.5)	C.V 0.36 (0.35)	S.E 16.21 (16.6)
Q_1* 327.0 (342.0)	Q_3* 535.50 (559.0)	D_1* 286.0 (282.0)	D_9* 687.20 (660.0)
t_3 0.163	t_4 0.149		

Note: The values in the parenthesis are for 85 years. *indicates values in 000-cusec.

Table-2
Parameters of P-III and LP-III Distributions (Kotri) (in 000's cusecs) 1901-1999

Distribution	Parameter	MOM	MLM	PWM
P-III	α	66.21	62.24	80.54
	β	5.81	6.498	4.08
	γ	76.53	47.55	123.49
LP-III	α	0.039	0.019	0.0194
	β	82.95	382.51	338.25
	γ	2.81	-0.55	-0.44

Table-3
Quantile Estimates and their S.E's (in parantheses)
by P-III and LP-III (Kotri) 1901-1999

Return Period T_r	Exceedence Probability P	Quanantile Magniude (000) cusecs)					
		Pearson III			Log-Pearson III		
		MOM	MLM	PWM	MOM	MLM	PWM
5	.20	575.47 (32.33)	575.57 (31.01)	574.802 (32.0)	573.40 (32.9)	573.41 (31.2)	572 (31.4)
10	.10	664.83 (42.45)	663.49 (41.13)	669.211 (42.8)	662.49 (42.4)	675.41 (41.23)	667.2 (41.8)
20	.05	746.05 (52.13)	743.04 (50.91)	756.405 (52.8)	743.69 (53.8)	774.65 (51.8)	750.4 (53.2)
25	.04	771.02 (55.16)	767.43 (53.71)	783.442 (55.9)	763.69 (56.0)	806.47 (53.9)	775 (55.1)
50	.02	845.87 (64.35)	840.41 (62.82)	865.085 (63.8)	843.73 (65.6)	905.75 (64.9)	861.1 (65.1)
75	.013	888.148 (66.59)	88.152 (64.19)	911.536 (66.1)	886.16 (70.5)	964.45 (68.2)	901.5 (69.3)
100	.01	917.57 (70.2)	910.11 (66.01)	944.026 (69.9)	915.73 (70.8)	1006.49 (69.9)	923.0 (70.2)

Table-4A
Ratios of Quantile Estimates (Kotri) 1901-1999

Ratio	Distribution/ Method	T_r					
		10	20	25	50	75	100
Q_{MOM} / Q_{MLM}	P-III	1.002	1.004	1.005	1.006	1.008	1.008
Q_{MLM} / Q_{MOM}	LP-III	1.02	1.04	1.05	1.07	1.09	1.09
Q_{PWM} / Q_{MLM}	P-III	1.01	1.02	1.02	1.03	1.03	1.04
Q_{PWM} / Q_{MOM}	LP-III	1.007	1.01	1.01	1.02	1.02	1.03
Q_{P-III} / Q_{LP-III}	MOM	1.003	1.004	1.005	1.002	1.002	1.002
Q_{P-III} / $Q_{LP(3)}$	MLM	0.98	0.96	0.95	0.93	0.91	0.90
Q_{P-III} / Q_{LP-III}	PWM	1.003	1.008	1.010	1.005	1.01	1.002

Table-4B
Ratios of S.E's of Quantile estimates (Kotri) 1901-1999

Ratio	Distribution/ Method	T_r					
		10	20	25	50	75	100
$S.E_{MOM}$ / $S.E_{MLM}$	P-III	1.03	1.02	1.04	1.03	1.05	1.06
$S.E_{MOM}$ / $S.E_{MLM}$	LP-III	1.03	0.02	1.04	1.02	1.01	1.01
$S.E_{P-III}$ / $S.E_{LP-III}$	MOM	0.99	0.99	0.98	0.97	0.97	0.99
$S.E_{P-III}$ / $S.E_{LP-III}$	MLM	0.997	0.98	0.996	0.97	0.96	0.94

Appendix

An Example of Fitting P-III Distribution by PWM to Annual Peak Flood Discharges Data at Kotri

a. Station Description

River: Indus
Barrage: Kotri
Period of records: 99 years
Data: See Fig. 1

b. Computational Procedure

Step 1:

$l_1 = 451.949$, $l_2 = 88.507$ $C_s = 0.830$, $C_v = 0.353$, $t_3 = 0.163$ and $t = 0.196$

Step 2:

Parameter estimation by PWM method.
Since $t_3 = 0.196$ is less than 1/3

$$t_m = 3 \;\; \pi \;\; t_3^2$$
$$= 3 \;\; (22/7) \;\; (0.163)^2 \; = \; 0.250$$

$$\hat{\beta} = \frac{\left(1 + 0.2906 \; t_m\right)}{\left(t_m + 0.1882 \; t_m^2 + 0.0442 \; t_m^3\right)} \overline{e}_{x_1}$$
$$= \frac{(1 + 0.2906 \times 0.250)}{(0.250 + 0.1882 \; (0.250)^2 + 0.0442 \; (0.250)^3}$$
$$= 4.078$$

$$\hat{\alpha} = \sqrt{\pi} \;\; l_2 \frac{\Gamma\left(\hat{\beta}\right)}{\Gamma\left(\hat{\beta} + 1/2\right)}$$

$$\Gamma\left(\hat{\beta}\right) = \Gamma(4.078) = 6.648$$

$$\Gamma\left(\hat{\beta} \; + \; 1/2\right) = \Gamma(4.078 + 1/2) = \Gamma(4.578) = 12.962$$

$$\hat{\alpha} = \sqrt{\pi} \; (88.507) \times \frac{6.48}{12.962}$$

$$\hat{\alpha} = 80.543$$

$\hat{\gamma} = l_1 - \hat{\alpha}\hat{\beta}$

$= 451.949 – (80.543)\ (4.078) = 123.495$

The fitted P-III Quantiles are obtained by

$$\hat{x}_T = \hat{\alpha}\ \hat{\beta} + \hat{\gamma} + K_T \sqrt{\hat{\alpha}^2 \hat{\beta}}$$

where

$$K_T = \frac{2}{C_s}\left[\left\{\frac{C_s}{6}\left(u - \frac{C_s}{6}\right) + 1\right\}^3 - 1\right],\ C_s > 0$$

and u is given in Table 7.

Step 3:

Computation of design flood estimates.

For T= 5, K_t = 0.755, $\hat{x}_5$ = 574.8022

For T= 50, K_T = 2.540, $\hat{x}_{50}$ = 865.085

For T = 100, K_T = 3.025, $\hat{x}_{100}$ = 944.0264

Table 7
Values of the standard normal variate

T (Years)	2	5	10	20	50	100
u_T	0	0.842	1.282	1.645	2.054	3.326

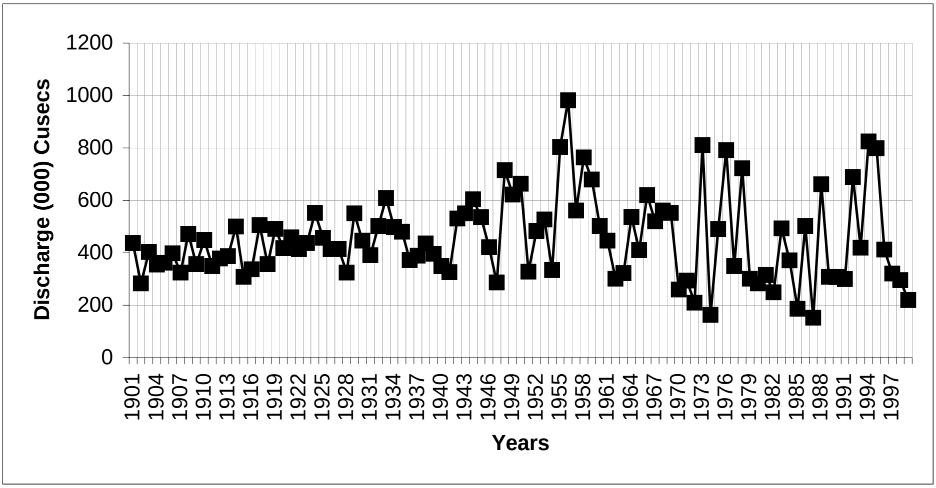

Fig. 1: Historigram of Peak Floods Discharge at Kotri Barrage (1901-1999)

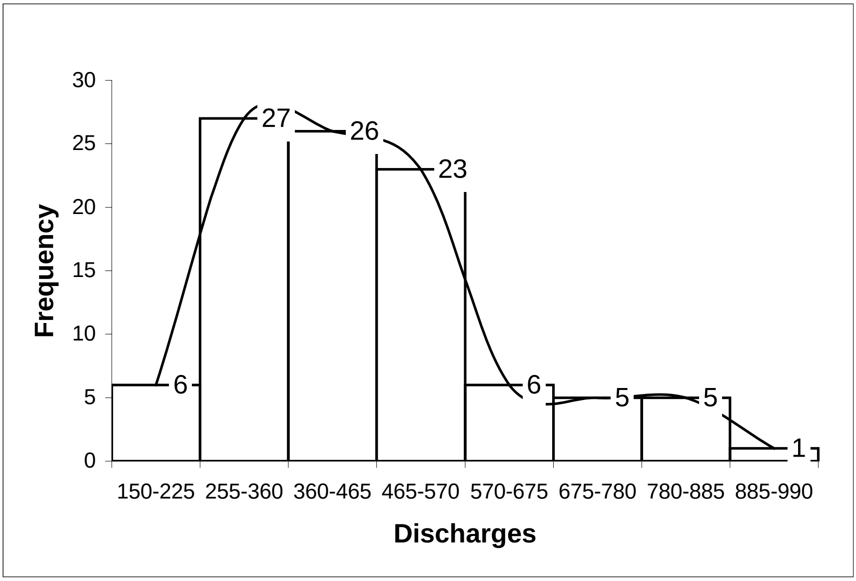

Fig. 2: Histogram & Frequency Curve of Flood Peaks at Kotri Barrage (1901-1999)

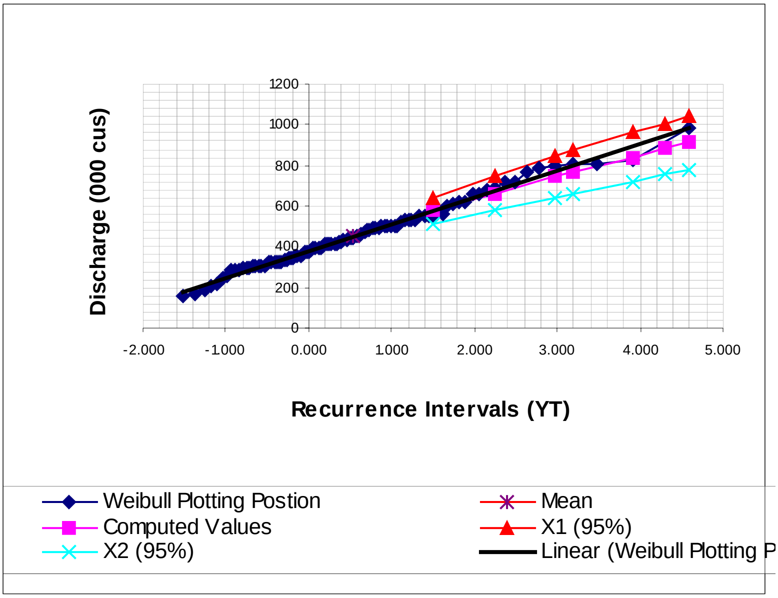

Fig. 3: Plot of P-III Distribution cure by MLM and Confidence Intervals of Flood Peaks of Indus River at Kotri Barrage (1901-1999)

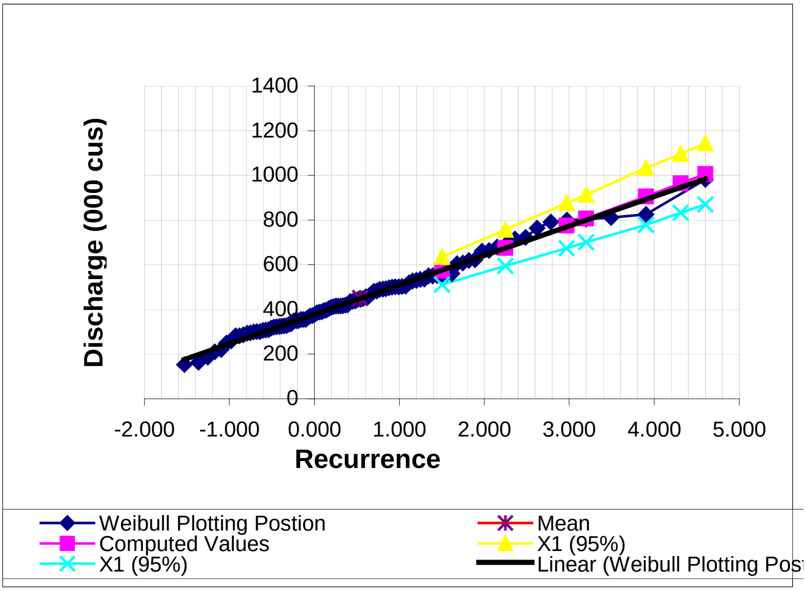

Fig. 4: Plot of LP-III Distribution curve by MLM and Confidence Intervals of Flood Peaks of Indus River at Kotri Barrage

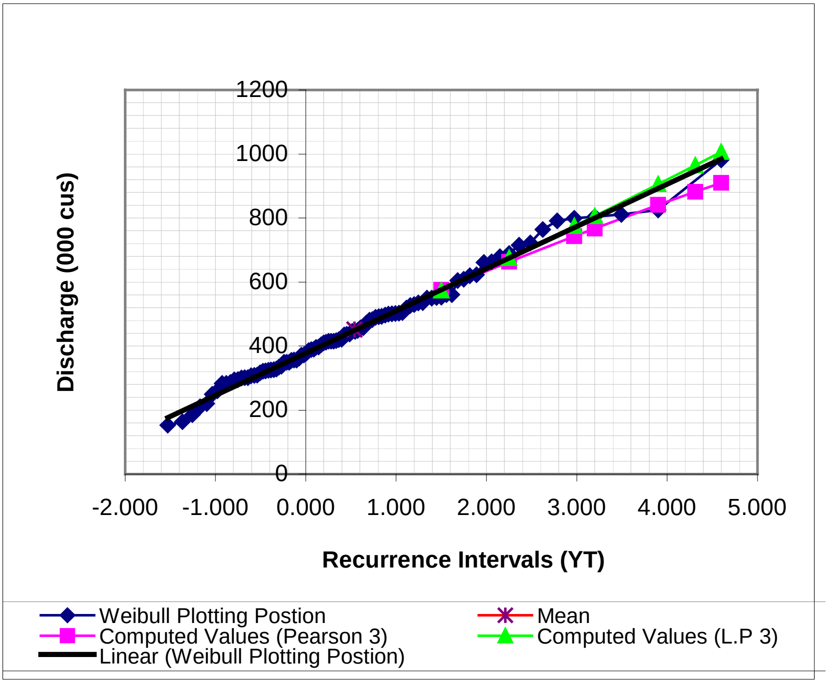

Fig. 5: Plot of P-III and LP-III Distribution Curves (by MLM) of Flood Peaks of Indus River at Kotri