Frequency Analysis of Seismic Data for Imminent Hazards

Faqir Muhammad
Department of Mathematics and Statistics,
Allama Iqbal Open University,
Islamabad, Pakistan 44000
aioufsd@yahoo.com

Gulshan Ara Majid
Department of Mathematics and Statistics,
Allama Iqbal Open University,
Islamabad, Pakistan 44000
gulshanaramajid@hotmail.com

Ali Daud
Department of Computer Science and Technology,
Tsinghua University
Beijing, China 100084
ali_msdb@hotmail.com

Abstract -**The seismic data are analyzed to find a suitable distribution for earthquake magnitude. Lognormal, Gumbel and Gamma distributions are implemented for earthquake aftershock magnitude using Pakistan's data from 8 Oct. 2005 to 17 Sep. 2006 and historical earthquake magnitude data between 1905 and 2006. Based on goodness of fit tests, it was found that Lognormal and Gumbel fit closely. Gumbel distribution is used for finding the return periods and hazard of earthquake. Experimental results show that it is able to assist more dependable estimation of seismic hazard, seismicity and future hazards.**

Keywords: Seismic hazards, Lognormal, Gamma, Gumbel distribution, Recurrence times

I. INTRODUCTION

Pakistan has experienced several damaging earthquakes over the last 100 years, but only three that must be characterized as national disasters, the Quetta earthquake in 1935, the Makran coast earthquake with tsunami generation in 1945 and the latest Kashmir earthquake on the 8th October 2005. The last earthquake improves the awareness about buildings with poor seismic resistance capacities.

In this study, Lognormal, Gamma and Gumbel models have been considered to describe earthquake magnitude data. Also these models were used to compute future hazard in the form of return periods and recurrence times. We can say that the solution provided by us is well justified and produced quite promising and functional results.

The rest of the paper is organized as follows. Section II provides our approach with explaining model estimation procedure. In section III, we discuss experimental results with details of models estimation and finding the return periods and section IV brings this paper to conclusions.

II. METHODOLOGY

The secondary data about earthquake aftershocks measured in Richter scale from 2005-2006 and historical earthquake data 1905-2006 of sub-continent collected by Pakistan Meteorology Department (PMD) Quetta and Peshawar Pakistan were used. The map of Pakitan available at PMD website gives different seismic zones of Pakistan. There were 2094 aftershocks from 08/10/2005 to 17/09/2006.

Also, the historical data between 1905 and 2006 had Magnitude from 3.8 to 8.6 Richter scale.

A. Model Estimation

The quantity of interest to be modeled is earthquake data which is positive so the skewed distributions will be suitable candidate. We tried Lognormal, Gamma and Gumbel distributions as models of earthquake magnitudes. The closely fitted distribution is used for prediction purposes as the results would assist more dependable estimation of seismic hazard and seismicity.

A considerable element of the statistician's exertion is to reach a realistic probabilistic model, by using the data to estimate the unknown parameters that characterize the model. Estimation gives a methodical and sensible way of deriving functional form of estimators. A number of studies have shown that MM and ML estimators are efficient for large samples, Hosking [1] but for small samples L-moments are better. As the number of values is large, the methods used for estimating parameters are maximum likelihood (ML), method of moments (MM) for Lognormal, Gumbel and Gamma distributions.

The Lognormal distribution has been used for earthquake recurrence intervals, Nishenko and Buland [2]. If *X* is a random variable with a normal distribution, then exp(*X*) has a log-normal distribution.

The log-normal distribution has probability density function:

$$f(x;\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{\frac{-(\ln x - u)^2}{2\sigma^2}} \qquad (1)$$

$$0 \le \mu, x \le \infty, \sigma \ge 0$$

Where, μ and σ are the mean and standard deviation of the variable's logarithm.

The M.L and M.M of Lognormal distribution are reviewed by Jhonson et al.[3].

Gamma distribution has been considered to describe the properties of Earthquake such as the spatial size of earthquake, time between earthquakes, and seismic moments by Kagan[4], Kagan [5]. The probability density function of Gamma distribution is:

$$f(x) = \frac{1}{\beta^{\alpha}\Gamma\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}}, \quad 0 \le x < \infty \qquad (2)$$

where,

$$\Gamma\alpha = \sqrt{\frac{2\pi}{\alpha}} e^{\alpha[\ln\alpha - g(\alpha)]},$$

$$g(\alpha) = 1 - \left(\frac{1}{12\alpha^2}\right) + \left(\frac{1}{360\alpha^4}\right) + \left(\frac{1}{1260\alpha^6}\right),$$

This distribution has no location parameter while β and α that are scale and shape parameters respectively. The M.L estimates are:

$$\hat{\alpha} = \frac{1}{4A}\left[1 + \sqrt{1 + \frac{4A}{3}}\right], \qquad (3)$$

here $A = \log_e \overline{X} - \left(\frac{\sum \log_e X}{n}\right)$, $\hat{\beta} = \frac{\overline{X}}{\hat{\alpha}}$,

The distribution of earthquake magnitude M, have long been assumed to follow Gumbel distribution, Gutenberg and Ritcher [6].

$$log(\mathrm{N}) = \mathrm{a} - \mathrm{bM}, \qquad (4)$$

where N is the number of earthquakes which exceeded magnitude M, and a & b are constants.

The theory of extreme values is widely used in Statistical model for determining the seismic characteristics of a region Epstein and Lomnitz [7], Yegulalp and Kuo [8] and Burton [9].

The Gumbel type I extreme value model whose distribution function is

$$F(x) = \exp\left[-\exp\left\{-\left(\frac{x-\mu}{\alpha}\right)\right\}\right], \qquad (5)$$

$$M < x < \infty$$

where M is the threshold extreme value of Aftershock magnitude (M_L).

If T is in years and the probability of Q being exceeded in a year is p. The return period is defined as the probability of exceedance in one year i.e. $T = p^{-1} = (1-F)^{-1}$, where F is the distribution function. The MM estimates of μ and α for Gumbel distributions are:

$$\hat{\alpha} = \frac{\sqrt{6}}{\pi}\hat{\sigma}, \quad \hat{\mu} = \overline{X} - 0.45\hat{\sigma}, \qquad (6)$$

Christopiet (1994) worked on the method of moments to fit the data for extreme value distribution i.e. Gumbel distribution. Bilim [10] found that any probability of occurrence of the earthquake of magnitude M within any D period is given by

$$R(M) = 1 - e^{-DN(M)}, \qquad (7)$$

where *N(M)* is the number of earthquakes of magnitude greater than M. These models have also been discussed by Utsu[11], Christopeit[12] and Ferra'es[13].

III. RESULTS AND DISCUSSION

The basic statistics for the aftershock magnitude and historical earthquake data in Richter of sub-continent are calculated and are presented in Table 1. It is evident that Aftershocks magnitude ranges from 2.5 to 7.5 and there was a variation of aftershocks with standard deviation 0.59. The Aftershocks magnitude followed a positive extreme with skewness 1.20. Histograms of selected aftershocks data along with normal curve are in Figure 1. This shows that distributions of aftershock is positively skewed i.e. curve is asymmetrical being stretched out to the right.

Table 1. a: Summary statistics of Aftershocks data and Percentiles

N	$\overline{X}$	S.D	β_1	SE Skew	β_2	SE Kurt	Min	Max
2094	3.6	0.59	1.2	0.05	1.55	0.11	2.5	7.5

Percentiles

10	25	30	50	60	70	75	80	90
3.0	3.1	3.2	3.4	3.6	3.9	4.0	4.2	4.4

Table 1. b: Summary Statistics of Historical Data and Percentiles

N	$\overline{X}$	S.D	β_1	SE Skew	β_2	SE Kurt	Min	Max
336	5.1	0.73	0.91	0.13	2.35	0.27	3.8	8.6

Percentiles

10	25	30	50	60	70	75	80	90
4.1	4.7	4.9	5.1	5.2	5.3	5.5	5.5	5.6

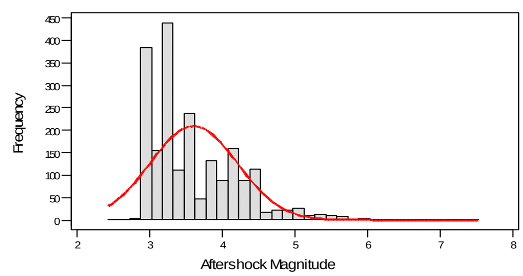

Figure 1. Histogram of aftershocks magnitude (M_L)

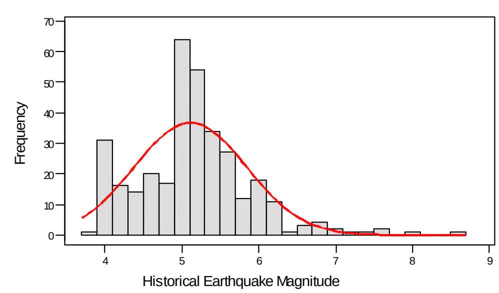

Figure 2. Histogram of historical magnitude (M_L)

Similarly historical earthquake magnitude ranges from 3.8 to 8.6 and there was a variation of earthquake magnitude with standard deviation 0.73. The earthquake magnitude followed a positive extreme with skewness 0.91 which is also clear from Figure 2. It shows that distribution of historical earthquake magnitude is positively skewed i.e. curve is asymmetrical being stretched out to the right.

A. ESTIMATION OF MODELS

In seismic hazard analysis studies, the main objective is to find the best fitted probability density function. As the sample size is large, two methods of estimating parameters of Lognormal, Gamma, and Gumbel distributions were used i.e ML and MM. The the fitted distributions were subjected to measures of goodness of fit such as sum of squares of errors(SSE) sum of absolute errors(SAE) and Kolmogorov Smirnov(KS). It came out that Gumble fits the data more closely and lognormal is the second.

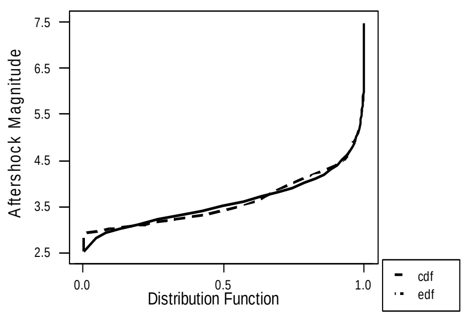

Figure 3. Aftershock earthquake magnitude and fitted lognormal distribution by M.M.

The plotted values for earthquake aftershock magnitude data of Pakistan and historical earthquake magnitude data are in Figures 3 and 4 respectively. It shows that Estimated CDF with ML fit both the data i.e. to aftershock and historical earthquake in a smooth manner and are closer with observed data. It can be concluded that ML is suitable method for estimating parameters of lognormal distribution.

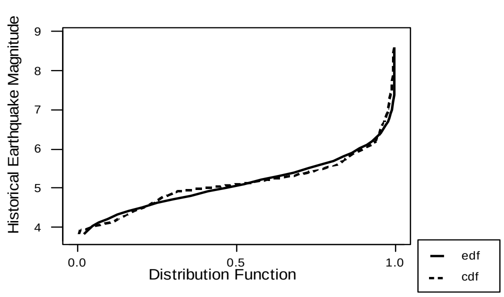

Figure 4. Historical Magnitude and Fitted Lognormal Distribution by ML.

The plotted values of Gumble fit with parameters estimated by MM for earthquake aftershock Fig magnitude data and historical earthquake magnitude data are shown in Figures 5 and 6 respectively. These figures show that parameters estimated by MM fit both the data It can be concluded that M.M is appropriate method for estimating parameters of Gumbel distribution.

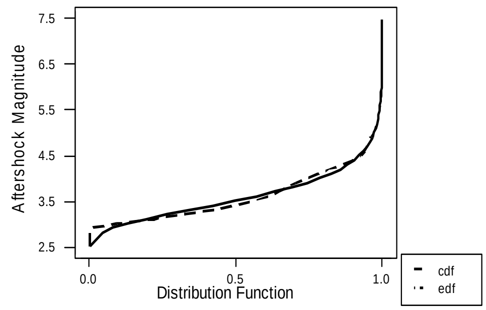

Figure 5. Aftershock earthquake magnitude and Fitted Gumbel distribution by MM.

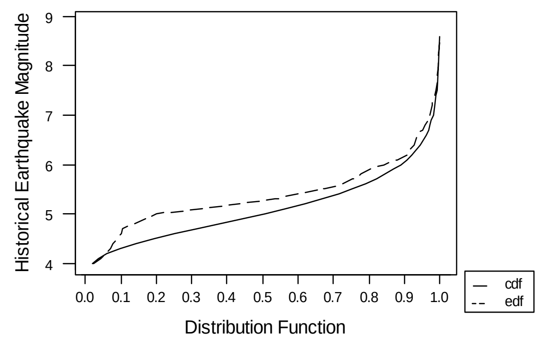

Figure 6. Historical magnitude and Fitted Gumbel distribution by MM.

Different goodness of fit criteria, such as sum of squares of errors, KS also indicated that lognormal and Gumbel fit these data well.

B. RETURN PERIODS

The return periods for historical earthquake magnitude are in Table 2 and graph is figure 7.

Table 2. Return Periods by Gumbel.

Magnitude (M_L)	**Return Periods in Years**
3.5	1.00
4.0	1.06
4.5	1.26
5.0	1.70
5.5	4.57
6.0	10.34
6.5	23.99
7.0	48.00
7.5	95.99
8.0	168.01
8.5	336.02

It is observed that the earthquake of magnitude 3.5 to 5.0 may occur within the duration of one or one and half years. The large earthquake of moderate 5.5 may occur within the period of 4.5 years and the earthquake of high magnitude of 6.0 and 6.5 may occur after every 10 and 23 or 24 years. But the major earthquake of 7.0 and 7.5 magnitudes may occur after the duration of 48 to 100 years respectively. The havoc earthquake of magnitude 8.0 and 8.5 may occur within 168 and 336 years respectively.

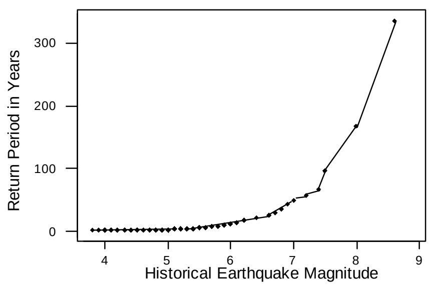

Figure 7. Historical magnitude versus return period.

The results of seismic hazard for predicted magnitude using Gumbel distribution are in table 3. These are the probabilities of occurrence by different magnitudes and years with 25, 50,…,150. The probability is 0.65 that an earthquake of 7.5 will occur in next 25 years and the probability is 0.99 that an earthquake of magnitude of 7.5 Richter will occur in next 125 years.

Table 3: Probability of earthquake occurrence using Gumbel Distribution

Gumbel Distribution						
Magnitude	**Probability of occurrence (%)**					
	25 years	**50 years**	**75 years**	**100 years**	**125 years**	**150 years**
4.0	92	99	100	100	100	100
4.5	99	100	100	100	100	100
5.0	100	100	100	100	100	100
5.5	100	100	100	100	100	100
6.0	99	99	100	100	100	100
6.5	92	99	100	100	100	100
7.0	80	96	99	100	100	100
7.5	65	88	96	98	99	100

IV. CONCLUSIONS

Lognormal and Gumbel models fit to both the datasets of earthquake aftershock magnitudes of (2005-2006) and historical earthquake magnitude (1905-2006). The return periods and probabilities of earthquake occurrence using Gumbel distribution indicate that a major earthquake of 7.0 and 7.5 magnitudes may occur after the duration of 48 to 100 years, respectively.

REFERENCES

[1] Hosking, J. R. M., (1990). L-moments: analysis and estimation of distributions using linear combinations or order statistics, J. Royal Statistical Soc. Ser. B., Vol. 52, 105-124.

[2] Nishenko, S., and Buland, R. (1987). A generic recurrence interval distribution for earthquake forecasting. *Bulletin of the Seismological Soc. of America*,.77,1382–1389.

[3] Jhonson, Norman L., Kotz, Samuel and Balakrishnan, N. (1995). Continuous Univariate Distribution Vol.2, Second Edition. *John Wiley and Sons Inc.*

[4] Kagan, Y.Y.,(1993).Statistics of characteristic earthquakes, *Bulletin Seismological Society of America*,Vol. 83, 7-24.

[5] Kagan Y. Y., (1997). Seismic Moment-frequency Relation for Shallow Earthquakes: Regional Comparison, *Journal of Geophysics Research*, Vol. 102, 2835-2852.

[6] Gutenberg B. and Ritcher C.F., (1954). Seismicity of Earth an Related Phenomena, 2nd. ed *Princeton University Press, Princeton*, New Jersey.

[7] Epstein B., and Lomnitz C., (1966). A model for the occurrence of large earthquakes: *Nature*, Vol. 211, 954-956.

[8] Yegulalp T.M. and Kuo J.J., (1974). Statistical prediction of the occurrence of maximum. Magnitude earthquakes, *Bull. Seism. Soc. Am.*, Vol. 64, 393-414.

[9] Burton P.W., (1979). Seismic risk in southern Europe through to India examined using Gumbel's third distribution of extreme values: *Geophys. J. R. Astr. Soc*., Vol. 59, 249-280.

[10] Bilim Funda, (2003). Investigation of seismic hazard in Cankiri, Turkey, using Gumbel's first asymptotic distribution of extreme values, *Journal of the Balkan Geophisical Society,* Vol. 6, 158-164.

[11] Utsu, T. (1984). Estimation of Parameters for Recurrence Models of the earthquakes. *Bull. Earthq. Res. Inst., Univ. Tokyo,* Vol. 59, 53-66.

[12] Christopeit, N., (1994). Estimating parameters of an extreme value distribution by the method of moments, Journal of Statistical Planning and Inference, Vol. 41, 173-186.

[13] Ferra`es, Sergio G. (2005). A probabilistic prediction of the next strong earthquake in the Acapulco-San Marcos segment, Mexico, Geofisica International, vol .44, 347-353.