Kashmir Journal of Science
https:\\kjs.org.pk

Kashmir Journal of Science (2022), 1(1): 60-67

Research Paper

Modelling the distribution of aftershocks data in Pakistan

Rijah Khan[1*] and Sayyad Khurshid [2]

[1]*School of Natural Sciences (SNS), National University of Sciences and Technology (NUST), Islamabad, Pakistan.*
[2]*Captain Hussain Khan Shaheed, Government Boys Post Graduate College Rawalakot, Azad Kashmir, Pakistan*
**Corresponding author's email:* rijahkhan2013@gmail.com

ARTICLE INFO

Article history:

Received: 22 September 2022
Accepted: 16 October 2022
Available online: 02-11-2022

Keywords:

Aftershocks data
Frechet distribution
Gamma distribution
Lognormal distribution
Weibull distribution

Abstract

In the present study an effort has been made to find out the best fitting distribution to explain the aftershocks behavior in Pakistan. For this purpose, two parameter Gamma, Weibull, Lognormal, Frechet and three parameter Generalized Extreme Value and Generalized Pareto distributions are fitted to aftershocks data and parameters for each distribution are estimated using the Maximum Likelihood method. The performance of the distributions is evaluated using Kolmogorove-Smirnove test. Finally, it is concluded that candidate distributions are found to be most appropriate distribution for describing the behavior of aftershocks data in Pakistan.

Introduction

An earthquake is a shaking of the ground caused by the sudden breaking and movement of large sections (tectonic plates) of the earth's rocky outermost crust. The edges of the tectonic plates are marked by faults (or fractures). Most earthquakes occur along the fault lines when the plates slide past each other or collide against each other. The shifting masses send out shock waves that may be powerful enough to alter the surface of the Earth, thrusting up cliffs and opening great cracks in the ground and cause great damage collapse of buildings and other manmade structures, broken power and gas lines (and the consequent fire), landslides, snow avalanches, tsunamis (giant sea waves) and volcanic eruptions.

Earthquakes are one of the most destructive of natural hazards. Earthquake occurs due to sudden release of energy in the Earth's crust that creates seismic waves. The impact of the event is most traumatic because it affects large area, occurs all on a sudden and unpredictable. They can cause large scale loss of life and property and disturbs essential services, communication and power, transport etc. They not only destroy villages, towns and cities but the outcome leads to destabilize the economic and social structure of the nation.

There are several types of earthquake waves including primary waves which are compressional and travel fastest, and secondary waves which are transverse, i.e., they cause the earth to vibrate perpendicularly to the direction of their motion.

The size of earthquake magnitude is determined by measuring the amplitude of the seismic waves recorded on the seismograph from the earthquake. These are put into the formula which converts them to a magnitude, which is a measure of the energy release by the earthquake.

Earthquake magnitude was traditionally measured on the Richter scale. It is often now calculated from seismic moment, which is proportional to the fault area multiplied by the average displacement on the fault.

The 2005 Kashmir earthquake, designates a major seismic activity with its epicenter in the Pakistan administered Kashmir. The earthquake occurred at 8:52:38 Pakistan standard time on October 8, 2005. It registered a debatable moment magnitude of 7.6 on the Richter scale.

There were several secondary earthquakes in the region. Primarily to the northwest of the original epicenter. A total of 147 aftershocks were registered in the first day after the quakes of which one had a magnitude of 6.2 Twenty-eight of these aftershocks occurred with magnitudes greater than the original quake. On October 19, a sequence of strong aftershocks, on with the magnitude of 5.8, occurred about 65KM (40miles) north north-west of Muzaffarabad. As of 27 October 2005, there have been more than 978 aftershocks with a magnitude of 4.0 and above that
continued to occur daily.

Figure1. 2005 Kashmir Earthquakes

In recent years many studies have been devoted on earthquake data by statistical analysis in different ways. e.g., Asanuma et al. (2014) applied a seismo statistical modelling method to micro seismic data collected at two geothermal field and investigated its feasibility for risk assessment. Moreover, Shcherbakov et al. (2012) studied statistical properties of the aftershock sequence of the Mw 7.1 in Darfield (Canterbury, New Zealand) earthquake. Singh et al. (2012) analyzed 3365 relocated aftershocks with magnitude of completeness (Mc)≥1.7 that occurred in the Kach Rift Basin (KRB) between August 2006 and December 2010. Rafi (2005) calculated the probability of occurrence of earthquake in Arabian Sea with different magnitude based on data records of Pakistan Meteorological Department and International since 1905-2002. Moreover, the earthquake related studies can be found in Cobanoglu and Alkaya (2011), Godinho (2007), Jones and Davies (1965), Kafka (2002), Kramer et al. (2002), Nasir et al. (2013), Naylor (2011), Rashke (2002), Shinozuka et al. (2011).

The purpose of this paper is to find the most appropriate distribution(s) for describing the behavior of aftershocks data. Therefore, we intend to compare Gamma, Weibull, Lognomal, Generalized Extreme value, Generalized Pareto and Frechet distributions based on Kolmogrov Simirnov (KS) test. Once a distribution function is assumed to be selected for study at hand, it remains to estimate its parameters from the sample data and to test the goodness of fit. Among the various estimation methods, Maximum Likelihood (ML) method is adopted in this study for parameters estimation of the candidate distributions.

Materials and Methods

About the Data

In the present study aftershocks (magnitude) data was obtained from earthquake Centre, Islamabad, Pakistan. The geographical position of station is shown in Figure 1.

Methodology

In order to describe the pattern of aftershocks data at a given station, it is spirited to recognize the distribution(s), which effectively fit the data. In this study two parameter Gamma, Lognormal, Weibull, Generalized Extreme value, Generalized Pareto and Frechet distributions were used to model the distribution of the aftershocks data along with Kolmogrove Simirnov (KS) test. The Probability Density Function (PDF) and Cumulative Distribution Function (CDF) for the mentioned models are given below.

Gamma distribution

The PDF and CDF of Gamma distribution having shape parameter (α), and scale parameter (β) are given by:

$$f(x;\alpha,\beta) = \frac{x^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} exp\left\{-\frac{x}{\beta}\right\}, x > 0, \alpha, \beta > 0\ ,$$

$F(x) = \frac{\Gamma\left(\frac{x}{\beta}\right)\alpha}{C\Gamma(\alpha)}$, where $\Gamma(x)$ is the incomplete Gamma function.

ML estimators for Gamma distribution.

The ML estimators of α and β are

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

where $= ln\bar{X} - \frac{\sum lnX}{n}$,

$$\hat{\beta} = \frac{\bar{X}}{\hat{\alpha}}\ .$$

Frechet distribution

The PDF and CDF of Frechet distribution are

$$f(x,\alpha,\beta) = \frac{\alpha}{\beta}\left(\frac{\beta}{x}\right)^{\alpha+1} exp\left\{-\left(\frac{\beta}{x}\right)^{\alpha}\right\}, x > 0, \alpha, \beta > 0$$

$$F(x) = exp\left[-\left(\frac{\beta}{x}\right)^{\alpha}\right], \;\; x > 0, \;\; \alpha, \beta > 0,$$

where the parameters α determines the shape of the distribution and β is the scale parameter.

ML estimators for Frechet distribution.

The log-likelihood function is

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

The partial derivatives of log-likelihood function are

$$\frac{\partial l}{\partial \alpha} = \frac{n}{\alpha} + nlog\beta - \sum_{i=1}^{n} log(x_i) - \sum_{i=1}^{n}\left(\frac{\beta}{x_i}\right)^{\alpha} log\left(\frac{\beta}{x_i}\right) = 0,$$

$$\frac{\partial l}{\partial \beta} = \frac{n\alpha}{\beta} - \left(\frac{\alpha}{\beta}\right)\sum_{i=1}^{n}\left\{\left(\frac{\beta}{x_i}\right)\right\}^{\alpha} = 0,$$

The solution of these equations does not yield closed form for the MLEs. However, a numerical method is applied for the simultaneous solution of these equations.

Lognormal distribution

The PDF and CDF of lognormal distribution are

$$f(x) = \frac{exp\left(-\frac{1}{2}\left(\frac{lnx-\mu}{\sigma}\right)^{2}\right)}{x\sigma\sqrt{2\pi}}, x > 0, \sigma > 0, -\infty < \mu < \infty$$

$$\boldsymbol{F(x) = \Phi\left(\frac{lnx-\mu}{\sigma}\right)}$$

Where Φ is Laplace integral.

ML estimators for Lognormal distribution

The ML estimators for μ and σ^2 are

$$\hat{\mu} = \frac{\sum_{i=1}^{n} lnx}{n}, \;\; \hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(lnx - \hat{\mu})^2}{n}$$

Weibull distribution

The PDF and CDF of Weibull distribution are

$$\mathrm{f(x)} = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left\{-\left(\frac{x}{\beta}\right)^{\alpha}\right\}, x > 0, \alpha, \beta > 0$$

$$F(x) = 1 - \exp\left\{-\left(\frac{x}{\beta}\right)^{\alpha}\right\}$$

ML estimators for Weibull distribution

The log-likelihood function is

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

The partial derivatives of log-likelihood function are

$$\frac{\partial l}{\partial \alpha} = \frac{n}{\alpha} - nlog(\beta) + \sum_{i=1}^{n} log(x_i) - \frac{\alpha}{\beta}\sum_{i=1}^{n}\left(\frac{x_i}{\beta}\right)^{\alpha-1} = 0,$$

$$\frac{\partial l}{\partial \beta} = -\frac{n\alpha}{\beta} + \left(\frac{\alpha}{\beta}\right)\sum_{i=1}^{n}\left\{\left(\frac{x_i}{\beta}\right)\right\}^{\alpha} = 0$$

These equations cannot be written in closed form. However, a numerical method is applied for the simultaneous solution of these equations.

Generalized Extreme Value (GEV) distribution

The GEV distribution with continuous location parameter (μ), continuous scale parameter (σ) and continuous shape parameter (γ) has the PDF and CDF are

$$f(x;\mu,\sigma,\gamma) = \frac{1}{\sigma}\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{\left(-\frac{1}{\gamma}\right)-1} exp\left\{-\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{\left(-\frac{1}{\gamma}\right)}\right\}$$

for $1+\gamma\left(\frac{x-\mu}{\sigma}\right) > 0,$

$$F(x;\mu,\sigma,\gamma) = exp\left\{-\left[1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right]^{\left(-\frac{1}{\gamma}\right)}\right\}$$

ML estimators for GEV distribution

The log-likelihood function for GEV distribution is

$$L = -nlog(\sigma) + \sum_{i=1}^{n}\left[\left(\frac{1}{\gamma}-1\right)\log(y_i) - (y_i)^{\frac{1}{\gamma}}\right]$$

Where $y_i = \left[1-\left(\frac{\gamma}{\sigma}\right)(x-\mu)\right]$. The Newwton Raphson method can be used to get the ML estimates of model parameters.

Generalized Pereto (GP) distribution

The PDF and CDF of GP distribution are

$$f(x) = \begin{cases}\frac{1}{\sigma}\left(1+k\frac{(x-\mu)}{\sigma}\right)^{-1-\frac{1}{k}} & k \neq 0 \\ \frac{1}{\sigma}exp\left(-\frac{(x-\mu)}{\sigma}\right) & k = 0\end{cases}$$

$$F(x) = \begin{cases}1-\left(1+k\frac{(x-\mu)}{\sigma}\right)^{-\frac{1}{k}} & k \neq 0 \\ 1-exp\left(-\frac{(x-\mu)}{\sigma}\right) & k = 0\end{cases}$$

ML estimators for GP distribution

The Newwton Raphson method can be used to get the ML estimates of model parameters of GP distribution.

Goodness of Fit Tests

To verify the goodness of fit for the fitted distributions in which X represent the aftershocks random variable and 'n' the sample size. Kolmogorov-Smirnov goodness of fit test is applied at 5% level of significance. The test is as follows:

Kolmogorov-Smirnov (KS) Test

KS test is a nonparametric alternative test of the Chi-square goodness of fit test and is used to confirm the sample under consideration selected from a reference or hypothesized distribution or to compare two samples come from identical distribution. The KS test statistics is calculated from the largest vertical difference in absolute value between the theoretical and the empirical cumulative distribution functions. Assume that we have a random sample from some distribution with CDF. The KS test statistic can be expressed as

Graphical analysis of aftershocks data

Histogram is constructed in Figure 1, which depicts that the distribution of aftershock is positively skewed i.e., the distribution curve is asymmetrical, being stretched out to the right due to high aftershocks. The other characteristic is that no aftershocks can be less than zero magnitude, thus the lower boundary of the distribution will be zero. This recommends that positively skewed distributions are appropriate candidates for aftershocks data. Moreover, to evaluate patterns of aftershocks data time series plot is built in Figure 2, where time ‘t’ is measured on the horizontal axis and the variable being observed is measured on the vertical axis.

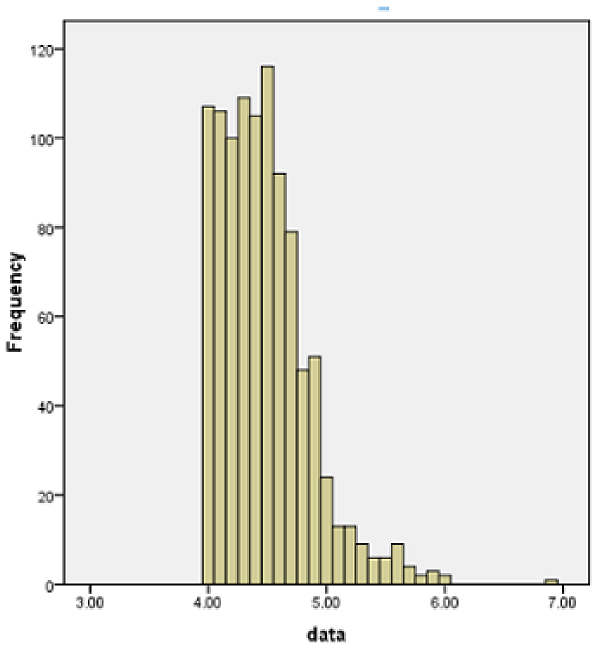

Figure1: Histogram of aftershock data

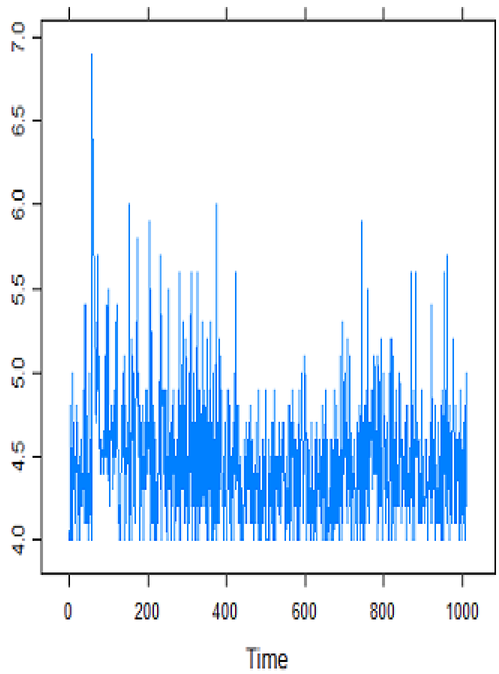

Figure 2: Time series plot of aftershock data

Basic analysis of aftershocks data

Summary statistics of aftershocks data is presented in Table 1. It reveals that aftershock ranges from 4 to 6.9. The coefficient of skewness is significant ($p<0.05$), which shows that the positively skewed distributions are appropriate for aftershocks data. Moreover, coefficient of kurtosis is $2.857<3$, indicates that the distribution of aftershocks data is platy kurtic. It is evident that positively skewed distribution(s) can be suitable candidates for aftershocks data.

Table 1: Summary Statistics for Aftershock data

n	Average	CV	SD	MIN.	MAX.	Skewness	Kurtosis
1001	4.4756	8.321	0.3779	4	6.9	1.271	2.857

Fitting Distributions

In this section GEV, Lognormal, Weibull and Gamma distributions are fitted to aftershocks data and parameters are estimated using ML method. Results are presented in Table 2 for comparison purposes.

Table 2: Parameter estimates for candidate distributions

Gamma distribution		Lognormal distribution		Weibull distribution		Frechet distribution	
α	**β**	**μ**	**σ**	**α**	**β**	**α**	**β**
141.2	0.0316	1.495	0.0807	14.186	4.6418	15.802	4.2984
GEV distribution				GP distribution			
γ	**σ**	**μ**		**k**	**σ**	**μ**	
0.0207	0.2852	4.3038		-0.3802	0.6623	3.9946	

KS is used to verify the goodness of fit for the fitted distributions and results are presented in Table 3. The p-value of KS is always far greater than 0.05 for all the distributions so all the distributions fit well to these data.

Table 3: P-values of KS test for candidate distributions

Gamma	Lognormal	Weibull	Frechet	GEV	GP
0.1016	0.0952	0.1650	0.0931	0.0831	0.0989

Conclusion

The intention of this paper is to compare different probability models namely, Gamma, Lognormal, Weibull, GEV, GP and Frechet for aftershocks data recorded in Pakistan. KS goodness of fit test is used to identify best fit at 5% level of significance. Based on these findings we concluded that all candidate distributions are most suitable for aftershocks data. Overall, this research also reveals that how applying parametric distributions with parameter estimates using aftershocks historical data. These results can be encouraging to step-up the information about the aftershocks history for particular area, local people, building contractors, scientists studying earthquake and storm water management planning.

References

Anagnos, T. and Kiremidjian, A. S. (1988). A Review of Earthquake Occurrence Models for Seismic Hazard Analysis. Journal of Stanford University, CA94305, USA, 3(1), 3-11.

Asanuma, H., Masaho, T. E. and Saeki, H. (2014). Seismostatistical Characterization of Earthquakes from Geothermal Reservoirs. University of Stanford California, 1-7.

Cobanoglu, I. and Alkaya, D. (2011). Seismic Risk Analysis of Denizli (Southwest Turkey) Region Using Different Statistical Models. International Journal of the Physical Sciences, 6(11), 2662-2670.

Godinho, J. (2007). Probabilistic Seismic Hazard Analysis. University of Patras, Greece, 1-49.

Jones, D. V. (1995). Forecasting Earthquakes and Earthquake Risk. International Journal of Forecasting, 11, 503-538.

Jones, D. V. and Davies, R. B. (1965). A Statistical Survey of Earthquake in the Main Seismic Region of New Zealand. Journal of Geology and Geophysics, 9(3), 251-284.

Kafka, A. L. (2002). Statistical Analysis of Hypothesis that Seismicity Delineates Areas Where Future Large Earthquakes are Likely to Occur in the Central and Eastern United State. Seismological Research Letters, 73(6), 992-1003.

Kramer, S. L., Jones, A. L. and Arduino, P. (2002). Estimation of Uncertainty in Geotechnical Properties for Performance-Based Earthquake Engineering. Journal University of Washington, 1-100.

Nasir, A., Lenhardt, W., Hintersberger, E. and Decker, K. (2013). Assessing the Completeness of Historical and Instrumental Earthquake Data in Austria and the Surrounding Areas. Austrian Journal of Earth Sciences, 106(1),90-102.

Naylor, M., Bell, A. F., Touati, S., Greenhough, J. and Main, I. G. (2011). Model Selection and Uncertainty in Earthquake Hazard Analysis. Journal Application of Statistics and Probability, 735-743.

Rafi, Z. (2005). Analysis of Seismicity in Arabian Sea Based on Statistical Model. Pakistan Journal of Meteorology. 2(4), 109-119.

Rashke, M. (2002). Modelling of Magnitude Distribution by the Generalized Truncated Exponential Distribution. International Seismological Centre, 1-12.

Shinozuka, M., Feng, M. Q., Kim, H., Uzawa, T. and Ueda, T. (2001). Statistical Analysis of Fragility Curves, Journal University of Southern California Los Angeles, CA 90089-2531.

Sheherbakov, R., Nguyen, M. and Quigley, M. (2012). Statistical Analysis of the 2010 Mw 7.1 Dar_eld Earthquake Aftershocks Sequence. Journal of Geology and Geophysics, 55(3), 305-311.

Singh, A. P., Mishra, O. P., Kumar, D., Kumar, S. and Yadav, R. B. (2012). Spatial Variation of Aftershocks Activity Across the Kachchh Rift Basin and its Seismotectonic Implication. Journal Indian Academy of Sciences, 121(2), 434-451.