Journal of Industrial Engineering International *Islamic Azad University, South Tehran Branch*
January 2008, Vol. 4, No. 6, 1-6

A new probability density function in earthquake occurrences

R. Sadeghian*

Ph.D. Student, Dep. of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran

G. R. Jalali-Naini

Assistant Professor, Dep. of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran

Abstract

Although knowing the time of the occurrence of the earthquakes is vital and helpful, unfortunately it is still unpredictable. By the way there is an urgent need to find a method to foresee this catastrophic event. There are a lot of methods for forecasting the time of earthquake occurrence. Another method for predicting that is to know probability density function of time interval between earthquakes. In this paper a new probability density function (PDF) for the time interval between earthquakes is found out. The parameters of the PDF will be estimated, and ultimately, the PDF will be tested by the earthquakes data about Iran.

Keywords: Forecasting; Probability Density Function (PDF); Distribution function; Earthquake

1. Introduction

Earthquakes are natural events that happen as tectonic stress or energy is released from faults. Faults are fractures on Earth's crust. Because of the interactions taking place beneath the crust, some energy is added to these faults. As the sum of these energies becomes unsustainable, the fault breaks open and hence the earthquake occurs [1,10].

With time, we may observe a repetitive pattern of increasing energy under the Earth's crust and its consecutive violet release from the fault. Once an earthquake has struck; the energy levels under the Earth's crust return to zero. This pattern of energy release may repeat itself time and time again.

Hence, time interval between two earthquakes is the same time interval between two releasing of energy from the fault. Thus knowing the time intervals between earthquake occurrences can be helpful for the prediction of next earthquake occurrence; even if, these time intervals are probabilities. Many researchers have studied well-known probability density functions for time intervals between two earthquake occurrences, recurrence time of an event, and waiting time to the next event. Some of them are pointed as follows:

A slip-predictable stochastic model was presented based on Markov renewal theory by Kiremidjian et al. Times between successive events are assumed to be Weibull-distributed with an increasing hazard rate. The model forecasts probabilities of earthquake occurrences conditioned to the time of occurrence of the last event along the Middle America Trench, Mexico [8].

Ferraes [4] believes that the determination of the probability distribution of recurrence times is the most important problem in the calculation of long-term conditional probabilities for the recurrence of large and great earthquakes. Hence, four different distributions were characterized by the property of maximum entropy, as possible laws for recurrence times of the largest earthquakes: uniform, exponential, Gaussian and log-normal in the west-northwestern zone of the Hellenic arc from 1791 to 1983.

The time distribution of main shocks, time intervals and seismotectonic segments of the Aegean area, was tested against the Poisson and negative binomial (Pascal) theoretical models [3].

Statistical analysis of seismicity data for the period 1907-1985 for events with magnitude $M \geq 6.0$ in the Hindu Kush and its vicinity were carried out to discuss the distribution characteristics of events and

* *Corresponding author. E-mail: sadeghian@iust.ac.ir*

seismicity variation using a stationary model of seismicity rates. The analysis reveals that the occurrences of earthquake follow Poisson's distribution [13].

Uniform, semi-Gaussian, Rayleigh, Weibull and lognormal statistical distributions of interval times and expected time to the next earthquake were compared by Sornette et al. [15]. Statistics of the recurrence times of great earthquakes at the Pacific subduction margins were made by Rikitake [12]. The probabilities of a great earthquake recurring in each zone are estimated on the basis of Weibull distribution analysis.

Refer to [16] for more references about the probability density functions of earthquake occurrences. These researches have continued during the last decade. Some of them are pointed as follows:

Ferraes [5] proposes a new method to estimate the time interval t for occurrence of a new large earthquake in Tokyo area using the probability distribution of recurrence times as described by the exponential, Weibull and Rayleihg probability densities form.

Kutch region of Gujrat is one of the most seismic prone regions of India. The probabilities of occurrence of large earthquake ($M \geq 6$ and $M \geq 5$) in a specified interval of time for different elapsed times were estimated on the basis of observed time-intervals between the large earthquakes ($M \geq 6$ and $M \geq 5$) using three probabilistic models, namely, Weibull, Gamma and Lognormal by Tripathi [17].

Also see [2,6,9] for more references about the probability density functions of earthquake occurrences.

In this paper, a new PDF for time intervals between earthquake occurrences is obtained. The Obtained PDF is different from the probability density functions used by the mentioned researchers. In the next sections, first, the method of obtaining the new PDF will be explained. Next, it is proved that the obtained function is a probability density function, and then the Mean and Variance of the PDF are obtained. In section of validation, both the history data of Iran area and a null hypothesis are used to prove that the PDF is valid. Finally, the proposed PDF is compared with other probability density functions.

2. Finding probability density function

Consider following symbols:

- y_i A quantity of cumulative energy on the time interval $[i-1,i]$
- U_n Total cumulative energy during the time interval $[0,n]$
- μ_0 Mean of cumulative energy on the time interval $[i-1,i]$
- σ_0^2 Variance of cumulative energy on the time interval $[i-1,i]$
- w Maximum level of energy endurable
- N A random variable that shows the number of times which energy is cumulated, and n is the value of this random variable.

Then,

$$U_n = \sum_{i=1}^{n} y_i \,. \tag{1}$$

Define: $G(n,w) = P(U_n > w)$.

If $n \to \infty$ then $\Delta t \to 0$, where Δt is the time interval between two cumulating energy. Therefore, based on CLT (Central Limitation Theorem) it can be written as follows:

$$P(N \geq n) = P(U_n \geq w) \,. \tag{2}$$

The above equation describe that the probability of accumulating energy more than w is equal to the probability that the number of accumulating energy is more a specific number as n, thus:

$$\left.\begin{array}{l} E(y_i) = \mu_0 \\ Var(y_i) = \sigma_0^2 \end{array}\right\} \Rightarrow \tag{3}$$

$$\Rightarrow \begin{cases} E(\sum_{i=1}^{n} y_i) = \sum_{i=1}^{n} E(y_i) = \sum_{i=1}^{n} \mu_0 = n\mu_0 \\ Var(\sum_{i=1}^{n} y_i) = \sum_{i=1}^{n} Var(y_i) = \sum_{i=1}^{n} \sigma_0^2 = n\sigma_0^2 \end{cases}$$

$$P(N \geq n) = P(U_n \geq w) \tag{4}$$

$$= P\left(\frac{\sum_{j=1}^{n} y_j - n\mu_0}{\sigma_0\sqrt{n}} > \frac{w - n\mu_0}{\sigma_0\sqrt{n}}\right) = 1 - \Phi\left(\frac{w - n\mu_0}{\sigma_0\sqrt{n}}\right),$$

where Φ is standard normal cumulative distribution function.

Therefore, corresponding to above equations, it can be written other equations as follows:

$$P(N<n)=1-(1-\Phi(\frac{w-n\mu_0}{\sigma_0\sqrt{n}}))=\Phi(\frac{w-n\mu_0}{\sigma_0\sqrt{n}})$$

$$=1-\Phi(-\frac{w-n\mu_0}{\sigma_0\sqrt{n}})=1-\Phi(\frac{n\mu_0-w}{\sigma_0\sqrt{n}})$$

$$\underset{n\ \text{is substituded by}\ t}{\overset{t:\ \text{non negative random variable}}{=\!=\!=\!=\!=\!=\!=\!=}}1-\Phi(\frac{t\mu_0-w}{\sigma_0\sqrt{t}})$$

$$=1-\Phi(\frac{w}{\sigma_0\sqrt{t}}(\frac{t}{w}\mu_0-1))$$

$$\underset{b=\frac{w^2}{\sigma_0^2}}{\overset{a=\frac{w}{\mu_0}}{=\!=\!=}}1-\Phi(\sqrt{\frac{b}{t}}(\frac{t}{a}-1))$$

$$\Rightarrow G(t)=1-\Phi(\sqrt{\frac{b}{t}}(\frac{t}{a}-1)). \quad (5)$$

Where $G(t)$ is survival function.

Therefore [7]:

$$-\frac{d(G(t))}{dt}=-\frac{d(1-\Phi(\sqrt{\frac{b}{t}}(\frac{t}{a}-1))}{dt} \quad (6)$$

$$\Rightarrow f(t)=\frac{1}{2\sqrt{2\pi}}\sqrt{\frac{b}{t}}(\frac{1}{a}+\frac{1}{t})e^{-\frac{1}{2}[\sqrt{\frac{b}{t}}(\frac{t}{a}-1)]^2},t\in[0,\infty)$$

Where $f(t)$ is probability density function for time interval between earthquakes, subject to t is a random variable as time between two successive earthquakes, also a,b are the parameters of the PDF.

Lemma 1. The function defined as above is a PDF.

Proof. Corresponding to the above definition, a,b,t are positive. So, $f(t)\geq 0$. Also,

$$I=\int_0^\infty f(t)dt=\frac{1}{2\sqrt{2\pi}}\int_0^\infty\sqrt{\frac{b}{t}}(\frac{1}{a}+\frac{1}{t})e^{-\frac{1}{2}[\sqrt{\frac{b}{t}}(\frac{t}{a}-1)]^2}dt. \quad (7)$$

Suppose, $u=\frac{1}{\sqrt{2}}\left(\sqrt{\frac{b}{t}}\left(\frac{t}{a}-1\right)\right)$; $t\in[0.+\infty)$

Therefore,

$$\left.\begin{aligned}\lim_{t\to 0^+}\frac{1}{\sqrt{2}}\left(\sqrt{\frac{b}{t}}\left(\frac{t}{a}-1\right)\right)=+\infty\\ \lim_{t\to 0^-}\frac{1}{\sqrt{2}}\left(\sqrt{\frac{b}{t}}\left(\frac{t}{a}-1\right)\right)=-\infty\end{aligned}\right\}\Rightarrow u\in(-\infty,+\infty)$$

$$u=\frac{1}{\sqrt{2}}\left(\sqrt{\frac{b}{t}}\left(\frac{t}{a}-1\right)\right)\Rightarrow du=\frac{1}{2\sqrt{2}}\left(\sqrt{\frac{b}{t}}\left(\frac{1}{t}+\frac{1}{a}\right)\right)dt$$

$$\Rightarrow I=\frac{1}{2\sqrt{2\pi}}\int_{-\infty}^{+\infty}2\sqrt{2}\,e^{-u^2}\,du=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-u^2}\,du$$

$$\xrightarrow{\text{it is clear that}\ \int_{-\infty}^{+\infty}e^{-u^2}du=\sqrt{\pi}\ [8]}I=\frac{1}{\sqrt{\pi}}\times\sqrt{\pi}=1. \quad (8)$$

The PDF is named as Jalali & Sadeghian PDF (JSPDF). ■

3. Cumulative distribution function of the JSPDF and its parameters

Consider Equation (6), it is clear that cumulative distribution function of the JSPDF (JSCDF) is as below [7]:

$$F(t)=\phi\left(\sqrt{\frac{b}{t}}\left(\frac{t}{a}-1\right)\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\sqrt{\frac{b}{t}}\left(\frac{t}{a}-1\right)}e^{-\frac{z^2}{2}}dz. \quad (9)$$

Lemma 2. In the JSPDF:

$$\mu_t=a\left(\frac{a}{2b}+1\right),\sigma_t^2=\frac{a^3}{b}\left(\frac{5a}{4b}+1\right). \quad (10)$$

Proof. Consider:

$$x=\left(\sqrt{\frac{b}{t}}\left(\frac{t}{a}-1\right)\right). \quad (11)$$

If X is a random variable with standard normal PDF, then t has JSPDF to notice the Equation (6). For obtaining mean and variance of JSPDF, obtaining $E(t), Var(t)=E(t^2)-E(t)^2$ suffices. Then consider:

$$x=\left(\sqrt{\frac{b}{t}}\left(\frac{t}{a}-1\right)\right)$$

$$\Rightarrow t=\frac{a^2x^2+2ab+ax\sqrt{a^2x^2+4ab}}{2b} \quad (12)$$

$$\Rightarrow \mu_t=E(t)=E\left(\frac{a^2x^2+2ab+ax\sqrt{a^2x^2+4ab}}{2b}\right)$$

$$=\frac{1}{2b}(a^2E(x^2)+E(2ab)+aE(x\sqrt{a^2x^2+4ab})).$$

It is clear that:

$$X\sim N(0,1)\Rightarrow X^2\sim\chi_1^2\Rightarrow E(X^2)=1.$$

Also, if $f(x)$ be an odd function, then $\int_{-a}^{a}f(x)dx=0$; for every arbitrary a [14].

So, because $x\sqrt{a^2x^2+4ab}$ is an odd function, then:

$$E(x\sqrt{a^2x^2+4ab})$$

$$=\int_{-\infty}^{+\infty}x\sqrt{a^2x^2+4ab}.\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx=0. \quad (13)$$

Similarly:

$$E(x^3\sqrt{a^2x^2+4ab})=0. \quad (14)$$

So,

$$\mu_t=\frac{1}{2b}[a^2.1+2ab+0]=a\left(\frac{a}{2b}+1\right). \quad (15)$$

It is clear that [11]:

$$X\sim N(0,1)\Rightarrow X^2\sim\chi_1^2\Rightarrow E(X^2)=1 \textit{ and}$$

$$Var(X^2)=2\Rightarrow Var(X^2)=E((X^2)^2)-E^2(X^2)$$

$$\Rightarrow 2=E(X^4)-1\Rightarrow E(X^4)=3. \quad (16)$$

Similarly:

$$E(t^2)=$$

$$=E\left(\frac{a^4x^4+4a^2b^2+a^2x^2(a^2x^2+4ab)+4a^3bx^2}{4b^2}\right)$$

$$+E\left(\frac{+2a^3x^3\sqrt{a^2x^2+4ab}+4a^2bx\sqrt{a^2x^2+4ab}}{4b^2}\right)$$

$$=\frac{1}{4b^2}[a^4E(x^4)+4a^2b^2+a^4E(x^4)$$

$$+4a^3bE(x^2)+4a^3bE(x^2)$$

$$+2a^3E(x^3\sqrt{a^2x^2+4ab})$$

$$+4a^2bE(x\sqrt{a^2x^2+4ab})]$$

$$=\frac{1}{4b^2}[3a^4+4a^2b^2+3a^4+4a^3b$$

$$+4a^3b+0+0]=\frac{6a^4+4a^2b^2+8a^3b}{4b^2}$$

$$\Rightarrow\sigma_t^2=E(t^2)-E^2(t)$$

$$=\frac{6a^4+4a^2b^2+8a^3b}{4b^2}-\frac{a^4+4a^2b^2+4a^3b}{4b^2}$$

$$=\frac{5a^4+4a^3b}{4b^2}=\frac{a^3}{b}\left(\frac{5a}{4b}+1\right). \blacksquare \quad (17)$$

4. Application and validation of JSPDF

In the previous section, in which JSPDF was obtained, the PDF had two parameters as a, b . Although JSPDF was employed for the data collected from just one fault, it can be applicable for the pieces of information from more than one fault. For example, information of earthquakes in all of Iran from 1973 to 2006 has been selected. The information has been collected from United States Geology Sciences Center site as below:

Http://www.neice.cr.usgs.gov/neis/epic/epic-rect.htm.

The information consisted of both foreshock data and aftershock data [10], so, first the data were filtered and the number of the data decreased from 4004 to 2606 [1].

Needed parameters have been computed as below:

$$\begin{cases} \mu = 4.65 \\ \sigma^2 = 30.53 \end{cases} \quad \text{(The time intervals are daily)}$$

$$\Rightarrow \begin{cases} \mu = a(\frac{a}{2b}+1) = 4.65 \\ \sigma^2 = \frac{a^3}{b}(\frac{5a}{4b}+1) = 30.53 \end{cases} \rightarrow \begin{cases} a = 2.649 \\ b = 1.757 \end{cases}$$

Consider:

$$\begin{cases} H_0 : \textit{The data has JSPdf} \\ H_1 : \textit{Otherwise} \end{cases}$$

Chi-square test is applied on the data and its p-value equals to 0.3242 and then H_0 is accepted when α level is equal to 0.05.

Therefore the Iran earthquakes data follow from JSPDF. Of course these data were of some different faults. The JSPDF acquires more satisfactory and reliable results, when employs data from just one fault.

Without considering the JSPDF, the Iran earthquakes data have been fitted by other known distributions. The fitness concluded that the data follow from Gamma density function with parameters $\alpha = 0.7682$, $\beta = 6.0528$. The Gamma density function was tested by chi square test and its p-value is equal to 0.15, so, H_0 is accepted when α level is equal to 0.05.

5. Conclusion

In this paper a probability density function for time interval of earthquakes data was found out, and the PDF named as JSPDF. Then, the parameters of JSPDF by μ, σ^2 have been obtained and The JSPDF was tested with the time interval of the Iran earthquakes data from 1973 to 2006, and the test was accepted by chi-square test. Next, the data were fitted with the Gamma density function. Therefore the time interval of every earthquake data follow from JSPDF, the Iran earthquakes data follow also from both JSPDF and Gamma PDF.

References

[1] Akasheh, B., 1999, *Basic Geophysics*. Farhang Peyk Press, Tehran, Iran.

[2] Corral, A., 2006, Dependence of earthquake recurrence times and independence of magnitudes on seismicity history. *Tectonophysics,* In Press.

[3] Dionysiou, D. D. and Papadopoulos, A. G., 1992, Poissonian and negative binomial modeling of earthquake time series in the Aegean area. *Physics of The Earth and Planetary Interiors*, 71(3-4), 154-165.

[4] Ferraes, G. S., 1991, A procedure to estimate the probability distribution of recurrence times in the west-northwestern Hellenic arc. *Earth and Planetary Science Letters,* 106(1-4), 181-190.

[5] Ferraes, G. S., 2003, The conditional probability of earthquake occurrence and the next large earthquake in Tokyo. *Journal of Seismology*, 7(2), 145-153.

[6] Fu, C. J. and Chang M. Y., 2003, On ordered series and later waiting time distributions in a sequence of Markov dependent multi-state trials. *Journal of Applied Probability,* 40(3), 623-642.

[7] Jalali Naini, S. G., 1997, *Stochastic Processes and its Applications in Biology, Agriculture, Management and Industries*. Statistics and Information Office Press, Tehran, Iran.

[8] Kiremidjian A. S. and Anagnos, T., 1984, Stochastic slip-predictable model for earthquake occurrences. *Bulletin of the Seismological Society of America,* 74(2), 739-755.

[9] Molchan, G., 2005, Interevent time distribution in seismicity: A theoretical approach. *Pure and Applied Geophysics,* 162(6-7), 1135-1150.

[10] Motohiko, H., Hassani, N. and Eslami, M. R., 2004, *Earthquake at the Album of Experience*.

Natural Critics Studies Center in Industry Press, Tehran, Second Edition, 10-17.

[11] Rezakhah, S., 1999, *Applied Probabilities and Statistics*. Poly Technique University Press, Tehran, Iran.

[12] Rikitake, T., 1976, Recurrence of great earthquakes at subduction zones. *Tectonophysics,* 35(4), 335-362.

[13] Shanker, D. and Singh, V. P., 1995, Earthquake distribution characteristics and seismicity rate in the Hindu Kush and its vicinity. *Acta Geophysica Polonica*, 43(4), 337-342.

[14] Silverman, R. A., 1988, *Modern Calculus and Analytic Geometry*, Scientific and Technical Press, Tehran, Iran.

[15] Sornette, D. and Knopoff, L., 1997, The paradox of the expected time until the next earthquake. *Bulletin of the Seismological Society of America,* 87(4).

[16] Tinti, S. and Mulargia, F., 1985, Application of the extreme value approaches to the apparent magnitude distribution of the earthquakes. *Pure and Applied Geophysics,* 123(2), 199-220.

[17] Tripathi, N. J., 2006, Probabilistic assessment of earthquake recurrence in the January 26, 2001 earthquake region of Gujrat, India. *Journal of Seismology,* 10(1), 119-130.