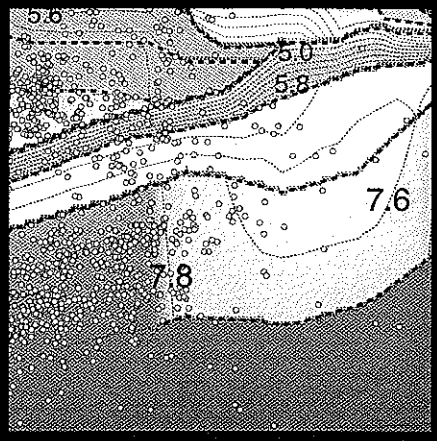
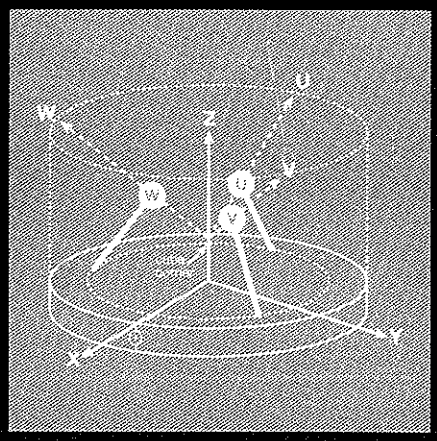
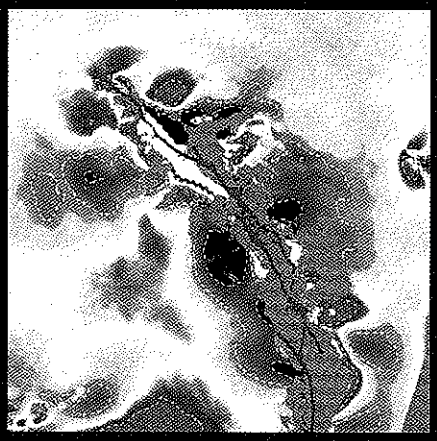


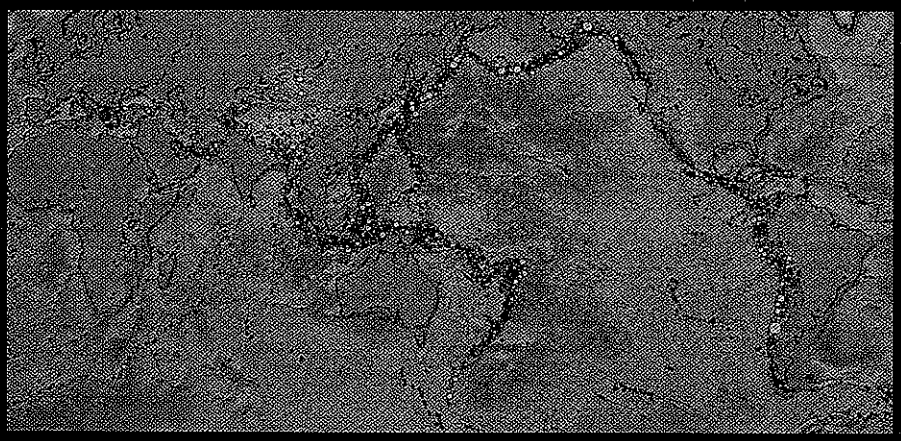


CD ENCLOSED

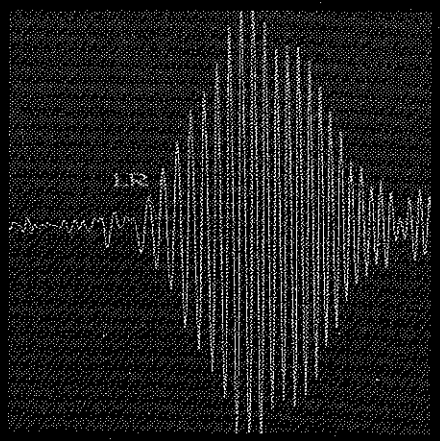
International Handbook of

EARTHQUAKE & ENGINEERING SEISMOLOGY





EDITED BY
William HK Lee
Hiroo Kanamori
Paul C Jennings
Carl Kisslinger



Part A

43

Statistical Features of Seismicity

Tokuji Utsu
The University of Tokyo, Tokyo, Japan

1. Introduction

Statistical studies of earthquake occurrences have frequently been carried out since the early years of seismology. To obtain reliable results from statistical analysis, a sufficient amount of high-quality data is necessary. The data are taken usually from earthquake catalogs, but many of the existing catalogs are inhomogeneous and incomplete. When we use seismicity data, special care must be taken to avoid or correct these defects in the catalogs.

In most statistical studies, earthquakes are represented by point events in a five-dimensional space–time–size continuum. In ordinary earthquake catalogs, the five coordinates are given as longitude and latitude of epicenter, focal depth, origin time, and magnitude. There are many other quantities which characterize an earthquake, fault-plane parameters (or more generally moment tensor components), stress drop, fault rupture length, rupture velocity, etc. Statistical studies involving these quantities are few. This is mainly because the complete data set on these is unavailable especially for small or old earthquakes.

The results of a statistical analysis must be tested for significance. Every method for a significance test is based on certain assumptions. When we use a method, we must remember the underlying assumptions.

It is often the case that we examine two or more data sets and choose the one for which the statistical property in question is most clearly recognized. The effect of such preferential selection of data set must be considered in the significance test. There is no common way of deciding how the significance level must be changed to compensate this effect. Numerical simulation may provide a solution in some cases.

2. Statistical Properties of Earthquake Sequences

2.1 Identification and Classification of Earthquake Sequences

Since the spatial and temporal clustering of earthquakes is the most prominent feature of most earthquake catalogs, we will first consider the properties of earthquake clusters, such as (foreshock–)mainshock–aftershock sequences and earthquake swarms.

There is no universally accepted exact definition of aftershocks, foreshocks, and earthquake swarms. We must give a working definition when we perform a statistical analysis on them. Figure 1a–d shows plots of occurrence rate versus time for typical earthquake sequences. The successive occurrence of mainshock–aftershock sequences of similar size (Fig. 1d) is also called an earthquake swarm, but it is apparently different from an ordinary swarm (Fig. 1c).

In statistical studies of seismicity, we sometimes use declustered catalogs, which contain only independent events (mainshocks and isolated earthquakes; the largest shock in a swarm is considered as the mainshock). Most algorithms for declustering (i.e., identification of earthquake sequences) use either space–time windows to include clustered events or space–time separations to link clustered events. For declustering or cluster identification algorithms, see Davis and Frohlich (1991a), Molchan and Dmitrieva (1992), and references in these papers. Studies using declustered catalogs or lists of earthquake clusters obtained by declustering are found in Utsu (1972a), Prozorov and Dziewonski (1982), Reasenberg (1985), Prozorov (1986), Keilis-Borok and Kossobokov (1990), Frohlich and Davis (1990), Davis and Frohlich (1991a,b), Ogata *et al.* (1995), and many others.

2.2 Aftershocks

2.2.1 Temporal Distribution of Aftershocks

Omori (1894) showed that the frequency of felt aftershocks per day, $n(t)$, following the 1891 Nobi, central Japan, earthquake ($M = 8.0$) decreased regularly with time according to the equation

$$n(t) = K(t + c)^{-1} \tag{1}$$

where K and c are constants and t is the time measured from the mainshock. The cumulative number of aftershocks, $N(t)$,

INTERNATIONAL HANDBOOK OF EARTHQUAKE AND ENGINEERING SEISMOLOGY, VOLUME 81A ISBN: 0-12-440652-1

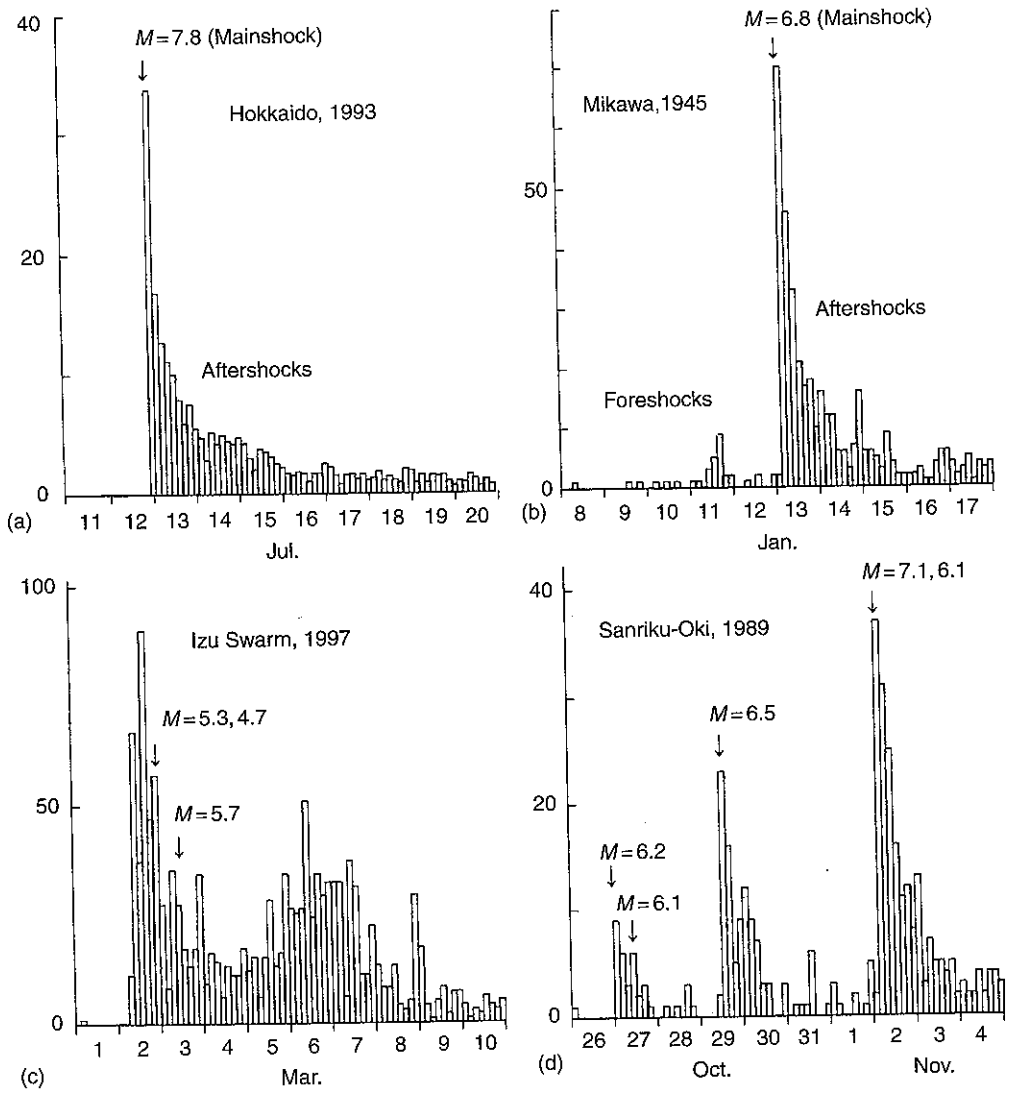


FIGURE 1 Examples of earthquake sequences. Temporal variations in the occurrence frequency per three hours are shown. (a) Mainshock–aftershock sequence; (b) foreshock–mainshock–aftershock sequence; (c) earthquake swarm; (d) successive occurrence of mainshock–aftershock sequences (earthquake swarm of the second kind).

plotted against $\log t$, tends to a straight line with time, because

$$N(t) = \int_0^t n(s)\,ds = K \ln(t/c + 1) \qquad (2)$$

Utsu (1961) plotted the data of $N(t)$ for 44 aftershock sequences against $\log t$, and showed that the slope of cumulative curve tends to decrease with time for most sequences. He reached the conclusion that Eq. (3) fits the data more closely and called it the modified Omori formula

$$n(t) = K(t + c)^{-p} \qquad (3)$$

where p is a constant somewhat larger than 1.0 (mostly 1.0–1.5). $N(t)$ for Eq. (3) is given by

$$N(t) = K\{c^{1-p} - (t + c)^{1-p}\}/(p - 1) \qquad (p \neq 1) \qquad (4)$$

If the origin times t_i $(i = 1, 2, \ldots)$ are available for all aftershocks of $M \geq M_z$ occurring in a certain period of time, we can compute the maximum likelihood estimates of K, p, and c (Ogata, 1983a). If the data contain the background seismicity of the region, it is recommended to use the equation $n(t) = K(t + c)^{-p} + B$ and estimate the background level B at the same time.

Computer programs for obtaining the maximum likelihood estimates of K, p, c, and B values and their standard errors for a given set of data are available in *IASPEI Software Library* Vol. 6 (Utsu and Ogata, 1997). This package contains the programs for the double sequence (the second sequence starts at $t = t_d$)

$$n(t) = K_1(t + c_1)^{-p_1} + K_2(t + t_d + c_2)^{-p_2} \quad (K_2 = 0 \text{ for } t < t_d) \tag{5}$$

and the triple sequence as well as some other formulas including the Weibull distribution

$$n(t) = K\alpha\beta t^{\beta-1} \exp(-\alpha t^{\beta}) \tag{6}$$

The stretched exponential function used by Kisslinger (1993) is the same as Eq. (6). Gross and Kisslinger (1994) tried more general equations.

The K value in the modified Omori relation depends strongly on the threshold magnitude M_z. The p value seems to be almost independent of M_z (e.g., Utsu, 1962). This is consistent with the stability of mean magnitude $\overline{M}$ (or b value) during an aftershock sequence (e.g., Lomnitz, 1966). The c value often shows strong dependence on M_z. This is mainly due to the deficiency of data. Small aftershocks occurring shortly after the mainshock tend to be missing due to overlapping of seismograms. One recent example supporting the modified Omori relation and the independence of p on M_z is shown in Figure 2 [see also Utsu *et al.* (1995) and Ogata (1999b)]. Nyffenegger and Frohlich (1998) investigated the factors that influence the maximum likelihood estimate of p value. Wiemer and Katsumata (1999) showed spatial variation of p and b values in the aftershock zones of four earthquakes (see also Utsu, 1962). Kisslinger (1996) presented a comprehensive review on aftershock phenomena.

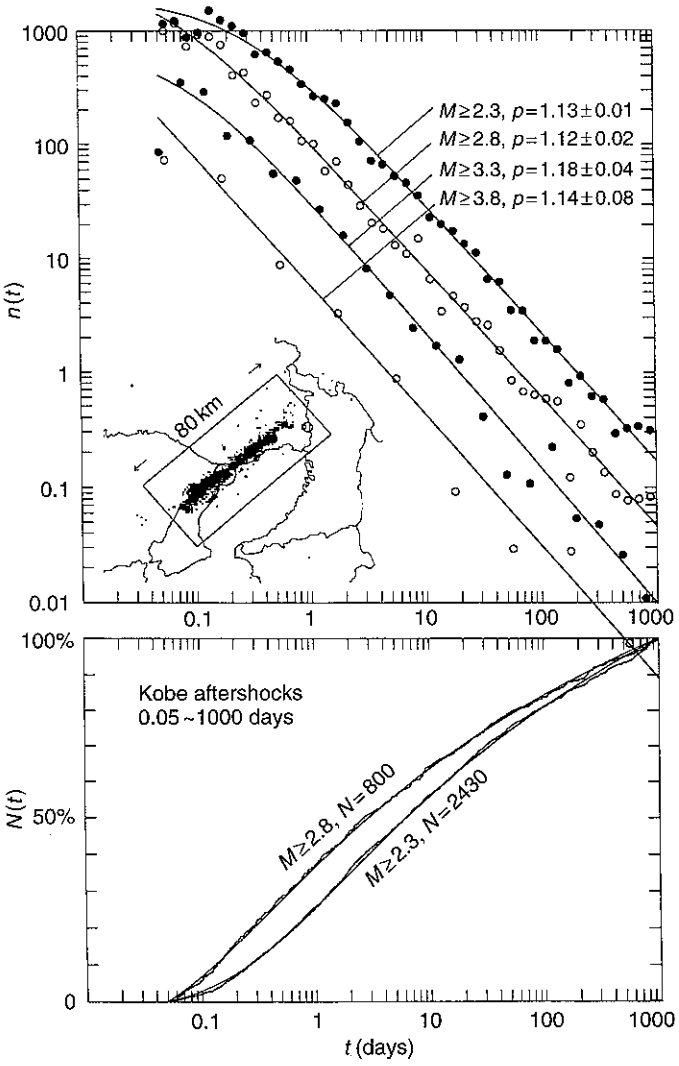


FIGURE 2 Modified Omori curves (top: rate, bottom: cumulative number) fitted to the aftershock sequences of the Hyogoken-Nanbu (Kobe) earthquake of 16 January 1995 (M_w 6.9). Data are taken from the Japan Meteorological Agency catalog. All earthquakes occurring in the rectangular area shown in the inserted map with focal depth less than 40 km are designated as aftershocks. The data and curves corresponding to the four (or two) different threshold magnitudes are shown. Note the independence of p value on the threshold magnitude.

2.2.2 Spatial Distribution of Aftershocks

It had been known before the introduction of earthquake magnitude that the aftershock epicenters scatter more widely for larger mainshocks. Utsu and Seki (1955) introduced the earthquake magnitude in this problem and obtained a relation $\log S = 1.02M_m - 4.01$, where S is the aftershock area in km^2 for a mainshock of magnitude M_m from the data of 40 mainshock–aftershock sequences in Japan ($6 \le M_m \le 8.5$). This equation can be written in a simpler form

$$\log S = M_m - 3.9 \tag{7}$$

If we consider a rectangular aftershock area with length L and width W and simply assume that $W = L/2$, Eq. (8) can be transformed to an equation proposed by Utsu (1961, 1969)

$$\log L = 0.5M_m - 1.8 \tag{8}$$

Although an aftershock zone roughly corresponds to the fault ruptured during the mainshock, precise studies indicate that aftershocks are concentrated near the margin of the fault area where the large displacement occurred (e.g., Hartzell and Heaton, 1986; Mendoza and Hartzell, 1988; Takeo, 1988). The relation between surface rupture length and M_m may have different coefficients (e.g., Tocher, 1958; Iida, 1965; Matsuda, 1975; Wells and Coppersmith, 1994), because the surface rupture does not always represent the entire fault rupture responsible for the earthquake.

There are many studies, which relate the size of aftershock zone to the mainshock magnitude. For example, Acharya (1979) obtained similar equations to Eq. (7) for five regions in

the Circum-Pacific zone, but the coefficient of M_m scatters from 0.78 to 1.22. Iio (1986) obtained

$$\log L = 0.43 M_m - 1.7 \quad (9)$$

for smaller earthquakes in Japan ($3.8 \leq M_m \leq 6.2$).

The aftershock zones sometimes show considerable expansion within days to years from the mainshock. For related studies, see Mogi (1968a) and Tajima and Kanamori (1985).

2.2.3 Aftershock Activity

The degree of aftershock activity relative to the mainshock size may be expressed by $\sum_{i=1}^{\infty} E_i/E_m$, where E_m and E_i denote the energy (or moment) of the mainshock and the ith largest aftershock, respectively. Since the energy (or moment) of the largest aftershock, E_1, is roughly proportional to the total energy (or moment), the aftershock activity is roughly represented by E_1/E_m, which corresponds to the magnitude difference between the mainshock and the largest aftershock

$$D_1 = M_m - M_1 \quad (10)$$

D_1 values for many mainshock–aftershock sequences have been investigated for its regional variation and dependence on M_m. Although $D_1 = 1.2$ is known as Båth's law (Richter, 1958), D_1 actually varies very widely from 0 to 3 or more.

The relation between D_1 and M_m has been studied by Utsu (1961) and several other workers. Utsu (1969) expressed $\tilde{D}_1$ (median of D_1 for mainshock magnitude M_m) as $\tilde{D}_1 = 5.0 - 0.5 M_m$ for $8.0 \geq M_m \geq 6.0$ and $\tilde{D}_1 = 2.0$ for $M_m \leq 6.0$ for shallow earthquakes in Japan. Most other workers gave a smaller average D_1 value for M_m less than about 6.5. For example, average D_1 for $M_m = 6.0$ is 0.67, 1.08, 1.14, and 1.05 according to Papazachos *et al.* (1967), Papazachos (1971), Båth (1977), and Kisslinger and Jones (1991), respectively. The difference between Utsu and most other workers may be caused by the difference in data-acquisition principle. Utsu considered the existence of earthquakes whose aftershocks were too small to be observed. Studies of regional variation of D_1 are found in Mogi (1967), Utsu (1969), Tsapanos (1990), Doser (1990), etc.

The aftershock activity has also been discussed by using the number of aftershocks above a certain magnitude level. For example, Singh and Suárez (1988) expressed the average number N of aftershocks of $m_b \geq 5$ accompanying large thrust earthquakes of moment magnitude M_w along the Circum-Pacific subduction zones. The relation is given by $\log N = M_w - 6.34$. Some systematic deviations from the average seem to depend on the degree of interplate coupling and complexities of the source. See Papazachos (1971), Lamoreaux *et al.* (1983), Yamanaka and Shimazaki (1990), and Davis and Frohlich (1991a) for regional trends.

Deep earthquakes are accompanied by few aftershocks (e.g., Wadati, 1931; Frohlich, 1987). Examples of aftershocks of deep earthquakes and depth variation of aftershock activity have been reported by Solov'ev and Solov'eva (1962), Prozorov and Dziewonski (1982), Pavlis and Hamburger (1991), Wiens *et al.* (1997), among others.

2.3 Foreshocks

Foreshocks are generally infrequent as compared with aftershocks. Foreshock activities are highly variable. Many large earthquakes of $M \geq 7.0$ are not preceded by foreshocks, even if the seismograph network is capable of detecting shocks of $M < 3.0$. Therefore the magnitude difference between the mainshock and the largest foreshock must be larger than 4 in these cases.

Temporal distributions of foreshocks are also quite variable. Foreshock sequences may show any pattern shown in Figure 1. Mogi (1985) classified the foreshock sequences into two types, C and D. In type C sequences, the activity increases gradually toward the mainshock. In type D sequences, the mainshock occurs after the foreshock activity dies down. Although no detailed statistics are available, type D seems to form a majority of foreshock sequences.

It is striking that, if we superpose the data of foreshock–mainshock time separations from many foreshock sequences, we obtain the power-law type distribution t^{-p} similar to the aftershock temporal pattern with the time direction reversed (Papazachos, 1974; Jones and Molnar, 1976, 1979; Davis and Frohlich, 1991b; Utsu, 1992; Ogata *et al.*, 1995; Maeda, 1999). Nevertheless, the great majority of individual foreshock sequences do not fit the power-law distribution. Only a limited number of foreshock sequences or precursory seismic activities show a regular increase of activity (power-law or exponential growth) toward the mainshock (e.g., Varnes, 1989; Jaumé and Sykes, 1999; Yamaoka *et al.*, 1999).

The discrimination of foreshock sequences from earthquake swarms is a difficult problem, because no easily recognizable differences in statistical or physical properties are found between them. For further discussion on this issue and other characteristics of foreshocks, including the relation to tectonics and fault types, see Mogi (1963), Jones and Molnar (1979), von Seggern *et al.* (1981), Jones (1984), Ogata *et al.* (1995, 1996), Abercrombie and Mori (1996), and Reasenberg (1999).

2.4 Earthquake Swarms

An earthquake swarm is a cluster of earthquakes in which there is no predominantly large single earthquake. Swarms are roughly divided into two types. The first type is the ordinary swarm in which the activity changes more or less irregularly (Fig. 1c) lasting a few hours to more than one year. Most swarms occurring in volcanic areas show this pattern, reflecting

the magmatic activity at depths. The second type is the successive occurrence of mainshock–aftershock sequences of similar size (Fig. 1d), which lasts several days to months.

3. Size Distribution of Earthquakes

3.1 Gutenberg–Richter Relation and its Equivalents

3.1.1 Power-Law Distribution

In general, smaller earthquakes are much more frequent than larger ones. Statistical studies of this property in late 19th century were not so successful because of the lack of adequate scale to measure the earthquake size. For historical review, see Utsu (1999), where the pioneering work of Wadati (1932) has been introduced. He adopted a power-law distribution for earthquake energy E

$$n(E) \propto E^{-w} \tag{11}$$

and estimated the value for exponent w as 1.7–2.1 or smaller by an indirect method. If the magnitude M is related to energy E (erg) by the equation of Gutenberg and Richter (1956)

$$\log E = 1.5M + 11.8 \tag{12}$$

Equation (11) is equivalent to the famous equation by Gutenberg and Richter (1944, 1949)

$$\log n(M) = a - bM \tag{13}$$

or

$$\log N(M) = A - bM, \qquad N(M) = \int_M^\infty n(M)dM \tag{14}$$

and $w = 1 + 2b/3$. Numerous studies indicated that Gutenberg and Richter's relation (hereafter called the G-R relation) is approximately valid in most cases and the value of b falls in the range 0.6–1.1. This corresponds to a w value of 1.40–1.73.

Ishimoto and Iida (1939) showed that the maximum amplitude A recorded at a seismograph station has a power-law distribution with exponent $m = 1.74$,

$$n(A) \propto A^{-m} \tag{15}$$

This is equivalent to the G-R relation with $b = m - 1$ (Asada *et al.*, 1951), if magnitude is defined by $M = \log A + f(\Delta)$, where $f(\Delta)$ is the calibrating function for epicentral distance Δ and the b value does not vary regionally. Most m values reported hitherto fall between 1.6 and 2.1.

Since the development of the concept of fractals, many phenomena exhibiting a power-law distribution have been treated as fractals. The G-R relation has often been interpreted in terms of the power-law (fractal) statistics of faults, cracks, and fragments (e.g., Takeuchi and Mizutani, 1968). Aki (1981) related the b value and the fractal dimension D as $D = 2b$. In these interpretations, it is implicitly assumed that each fault (or crack, or fragment) corresponds to one earthquake, i.e., the probability of earthquake occurrence is independent of the fault size. This assumption does not seem so natural.

3.1.2 Estimation of *b* Value

Under the assumption that the magnitude data are random samples from a population obeying the G-R relation, the method of moments (Utsu, 1965) and the method of maximum likelihood (Aki, 1965) both yield the solution

$$b = \log e/(\overline{M} - M_z) \tag{16}$$

where $\overline{M}$ is the mean magnitude of earthquakes of $M \geq M_z$ and $\log e = 0.434294$. M_z is the threshold magnitude above which the data should be complete. The standard error of the maximum likelihood estimate of b is approximately $b/\sqrt{N}$ for large N (the number of earthquakes of $M \geq M_z$). The smallness of error does not necessarily indicate the goodness of fit of the G-R relation to the data.

Since Eq. (16) is obtained for continuous exponential distribution, care must be taken when we use discrete (rounded) magnitude values. If the magnitudes are given at intervals of 0.1 as in most catalogs, and if we use the data with $M \geq 4.0$ for example, we must put $M_z = 3.95$ in Eq. (16), because $M = 4.0$ means $4.05 > M \geq 3.95$. If we use $M_z = 4.0$ in this case, the error in b value may easily reach several percent. There is another kind of error due to rounding (e.g., Tinti and Mulargia, 1987; Vere-Jones, 1989). This error is not large (usually less than 1%) if the magnitudes are given to the nearest 0.1 magnitude unit or less.

3.1.3 Temporal and Spatial Variation of *b* Value

Numerous papers have been published dealing with the spatial and temporal variation (or stability) of b value (e.g., Utsu, 1971; Li *et al.*, 1983; Imoto, 1987; Jin and Aki, 1989; Ogata and Katsura, 1993). For example, temporal changes in b value for background seismicity prior to large earthquakes have occasionally been reported (e.g., Smith, 1986, 1998; Trifu and Radulian, 1991; Imoto, 1991). Smaller b values for foreshocks than those for aftershocks and background seismicity have also been reported (e.g., Suyehiro *et al.*, 1964; Berg, 1968; Wu *et al.*, 1976; Molchan and Dimitrieva, 1990; Molchan *et al.*, 1999). However, we notice that many foreshock sequences have normal b values, and some earthquake swarms with small b values are not followed by future large earthquakes.

Spatial variations of b value have been studied for various regions of the world and for various focal depth ranges (Ogata *et al.*, 1991; Frohlich and Davis, 1993; Okal and Kirby, 1995; Wiemer and Wyss, 1997; Molchan *et al.*, 1997; Wiemer *et al.*, 1998, and many others). However, universally recognizable

regularities are relatively few. Some authors have argued against the regional variation of size distribution (e.g., Kagan, 1997, 1999).

To test the significance of the difference in b values between two earthquake groups, a method based on the F-distribution has been used (Utsu, 1966). A method using Akaike Information Criterion (AIC) (see Chapter 82 by Vere-Jones and Ogata) is simpler (Utsu, 1992, 1999). Of course these tests assume the G-R relation. If this assumption is not valid, it is possible that the two groups have nearly equal b values calculated from Eq. (16) but the size distributions are quite different.

3.2 Modified Equations

Although the G-R relation fits the data fairly well in many cases, significant deviation from it has often reported. The $\log n(M)$ (or $\log N(M)$) versus M plots for some data sets show considerable curvature, though the G-R relation predicts a straight line with slope of $-b$.

Various modifications of the G-R relation have been proposed to represent such curved distributions (see a review by Utsu, 1999). The simplest one is the truncated G-R relation, i.e., $\log n(M) = a - bM$ for $M \le M_{\max}$ and $n(M) = 0$ for $M > M_{\max}$. The maximum likelihood estimate of b in this case is somewhat different from that calculated from Eq. (16) (e.g., Page, 1968). If the truncated G-R relation is valid, the regional variation in b value calculated from Eq. (16) may be caused by the regional difference in $M_{\max}$ rather than the b slope in the range below $M_{\max}$.

The power-law distribution of energy or moment tapered by an exponential function used by Kagan (1991a, 1993, 1997) has the form (Gamma distribution)

$$n(E) = CE^{-w}\exp(-E/E_{\max}) \tag{17}$$

This is equivalent to

$$\log n(M) = a - bM - k10^{1.5M} \tag{18}$$

where $b = 1.5\,(w-1)$ and $k = 10^{-1.5M_{\max}}\log e$ ($M_{\max}$ is the magnitude corresponding to $E_{\max}$). This is a generalized form of an equation for a branching model (Otsuka, 1972) derived by Saito *et al.* (1973), Vere-Jones (1976), and Maruyama (1978), in which b takes a value of 0.75.

Another modification is the power-law distribution tapered by a logarithmic function

$$n(E) = CE^{-w}\ln(E_{\max}/E), \quad (n(E) = 0 \text{ for } E \ge E_{\max}) \tag{19}$$

which is equivalent to the equation proposed by Utsu (1971)

$$\log n(M) = a - bM + \log(M_{\max} - M) \tag{20}$$

It is possible to obtain the maximum likelihood estimates of the parameters of Eqs. (17)–(20) and some other equations including those proposed by Lomnitz-Adler and Lomnitz (1979), Makjanić (1980), Anderson and Luco (1983), and Seino *et al.* (1989) together with the AIC values for a given set of magnitude data. Computer programs are available in *IASPEI Software Library* Vol. 6 (Utsu and Ogata, 1997). We can compare AIC values for different equations and find the best fitting equation. For model selection criteria using AIC, see Chapter 82 by Vere-Jones and Ogata. Utsu (1999) provides actual examples.

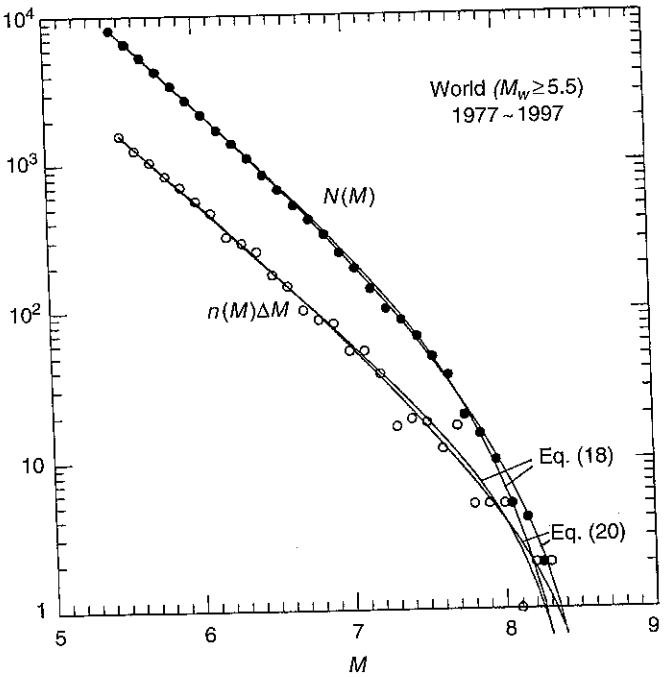


FIGURE 3 Frequency–magnitude distribution for earthquakes in the world for the years 1977–1997. The magnitude used is the moment magnitude (M_w) converted from seismic moments given in the Harvard University CMT catalog. Open and solid circles denote the frequency per 0.1 magnitude unit and cumulative frequency, respectively. Curves for Eqs. (18) and (20) fitted to the data are shown.

Figure 3 shows the distribution of M_w for recent earthquakes in the world. The curves represent Eqs. (18) and (20) fitted to the data. If the complete data in the 1950s and 1960s were available, the distribution would be closer to the G-R relation (straight line), since several great earthquakes of $M_w = 8.7$–9.5 occurred during 1952–1965 (Kanamori, 1977).

4. Temporal Distribution of Earthquakes

4.1 Stationary Poisson Process

The simplest reference model for the distribution of earthquakes in time is the stationary Poisson process, in which all events occur independently and uniformly in time. This process