Extreme value statistics and thermodynamics of earthquakes: aftershock sequences

Bernard H. Lavenda ([1])([2]) **and Elvio Cipollone** ([2])
([1]) *Università di Camerino (MC), Italy*
([2]) *National Centre for Research on Thermodynamics, ENEA-CR Casaccia, S. Maria di Galeria, Roma, Italy*

Abstract
The Gutenberg-Richter magnitude-frequency law takes into account the minimum detectable magnitude, and treats aftershocks as if they were independent and identically distributed random events. A new magnitude-frequency relation is proposed which takes into account the magnitude of the main shock, and the degree to which aftershocks depend on the main shock makes them appear clustered. In certain cases, there can be two branches in the order-statistics of aftershock sequences: for energies below threshold, the Pareto law applies and the asymptotic distribution of magnitude is the double-exponential distribution, while energies above threshold follow a one-parameter beta distribution, whose exponent is the cluster dimension, and the asymptotic Gompertz distribution predicts a maximum magnitude. The 1957 Aleutian Islands aftershock sequence exemplifies such dual behavior. A thermodynamics of aftershocks is constructed on the analogy between the non-conservation of the number of aftershocks and that of the particle number in degenerate gases.

Key words *Pareto and beta power laws – cluster dimension – frequency-magnitude regression laws – order-statistics – independence and clustering – thermodynamics of aftershocks*

1. Main shocks *versus* aftershocks

In many respects, aftershocks behave in the same manner as main shocks. Unlike most physical phenomena which decay exponentially, the decay of aftershock activity is hyperbolic in time. This was discovered over a century ago by Omori (1894), and it implies that aftershocks are basically a nonstationary process. The probability of occurrence of an aftershock decays in proportion to the time elapsed since the main shock. Moreover, it is generally accepted that the same frequency-magnitude, or Gutenberg-Richter (GR), law applies to aftershock sequences as it does to main shocks (Ranalli, 1969; Evison, 1999).

The GR law takes into consideration the minimum detectable magnitude, due in part to the finiteness of seismological networks. Whether there exists, or not, a maximum magnitude, which would be applicable to all regions, is still an open question (Knopoff and Kagan, 1977; Lomnitz-Adler and Lomnitz, 1978). In terms of energy, rather than magnitude, the GR law is a Pareto distribution. The Pareto distribution is a truncated distribution, applying, say, to salaries above a given level, which is analogous to a minimum detectable energy of an aftershock. The Pareto distribution has no right endpoint, and consequently, there is no maximum salary, or «maximum magnitude». Seen in these terms, the GR law asserts that there are more poor people than rich ones (Gumbel, 1958). While it

Mailing address: Prof. Bernard H. Lavenda, Università degli Studi di Camerino, Casella Postale 53, 62032 Camerino (MC), Italy; e-mail: lavenda@camserv.unicam.it

is true that the Pareto distribution does not possess a finite mean when its exponent is less than one, it cannot be used as an argument in favor of a maximum magnitude (Knopoff and Kagan, 1977), since the existence of a mean value of the energy has nothing to do with the finiteness of the total amount of energy released in an earthquake. In fact, the sum of all incomes is not even invariant for the maximum likelihood estimate of the Pareto exponent (Gumbel, 1958).

Therefore, if aftershocks follow the same GR relation as main shocks, they will be insensitive to the magnitude of the main shock. Clearly, the magnitude of the main shock is an upper limit for the magnitude of any aftershock in the sequence. This appears to be confirmed by the observation that the difference in instrumental magnitudes of the main shock and the largest aftershock in a sequence is roughly 1.2, independent of the absolute magnitude of the shock, or the particular nature of the aftershock sequence. This has come to be known as Båth's law (Richter, 1958; Båth, 1965). Although a relation between the magnitudes of the main shock and the largest aftershock is not to be unexpected, it appears to be independent of the number of distinct aftershocks in the sequence, or what constitutes the size of the sample. It is well-known from the testing of materials that their mean strength depends on the size of the sample that is being tested (Epstein, 1948). Just as the size of the specimen affects the distribution of strengths, so too should the mean magnitudes for the largest and smallest aftershocks vary as the number of aftershocks in the sequence. A dependence on the number of aftershocks has been proposed for the mode (Kurimoto, 1959), and the mean (Vere-Jones, 1969), of the magnitude of the main shock. Specifically, it has been suggested that the expected value of the magnitude of the main shock should vary as the logarithm of the number of aftershocks (Vere-Jones, 1969). The decrease of the mode as the logarithm of the number of shocks indicates that the individual shocks should be distributed according to a negative exponential distribution. Just as the strength of a specimen decreases as the specimen size increases, so too should the mean magnitude of an aftershock decrease as the number of aftershocks increases. The manner by which it decreases with the number of aftershocks will provide an indication of how the magnitudes are distributed.

If there is a relation between the cumulative frequency of aftershocks and the magnitude of the main shock it cannot be given by the GR law because it is independent of the magnitude of the main shock. With only a finite energy available, the length of the life span will be determined by the gradual depletion of energy from the earthquake zone. If the GR law applies to aftershocks then they are impervious to the energy of the main shock, which cannot generally be the case. Different distributions can be distinguished according to their «hazard rates», or the probability that a device will fail in a given interval. If the system is prone to initial failure, the hazard rate is a decreasing function of its argument. The Pareto distribution is an example. Rather, if the cause of failure is by chance, then the hazard rate remains constant. The exponential distribution is typical of such processes. Finally, if the cause of failure is wear out, then the hazard function is increasing. The Weibull distribution for the smallest value has an increasing hazard rate for an exponent greater than one. This distribution has a right endpoint which is characteristic of oldest age, or «maximum magnitude».

The power laws which characterize aftershock sequences are given as functions of energy, time and location:

Energy or magnitude

1) Aftershocks, like main shocks, supposedly follow Pareto's law in terms of energy, ε,

$$\bar{F}(\varepsilon) = \left(\frac{\varepsilon_\star}{\varepsilon}\right)^{\rho} \tag{1.1}$$

where $\bar{F}$ «survivor» function, or the tail of the distribution. The minimum detectable energy of the sequence is $\varepsilon_\star$ and $\rho > 0$. In terms of the magnitude m of an aftershock, (1.1) becomes the GR law

$$\log \boldsymbol{Nr}(> m) = a - bm \tag{1.2}$$

where $\boldsymbol{Nr}\,(> m)$ is the number of aftershocks having a magnitude greater than m. The coefficients in the GR law are

$$a = \log n + bm_{\star} \qquad (1.3)$$

and $b > 0$, where n is the total number of aftershocks in the sequence, and $m_{\star}$ (if $\neq 0$) is the minimum detectable magnitude in the sequence. Energy and magnitude are related by

$$\log \varepsilon = \alpha + \beta m \qquad (1.4)$$

where the constant α depends on the unit of energy chosen, and $\beta = 1.5$, according to Gutenberg (1956). Calculating the energy in ergs, $\alpha = 11.8$ (Utsu, 1969-1970). The exponent in the Pareto tail distribution is therefore $\rho = b/\beta$.

2) Magnitude stability in time. A mean magnitude is stationary in time, and fluctuations occur about this value with no appreciable decay in time for periods of up to 100 days. This was first proposed by Lomnitz (1966). However, it does not necessarily imply that the magnitude distribution is stationary; rather, it appears that the magnitudes are being selected from a stationary sample distribution at a rate which varies in time (Lomnitz, 1966). In other words, we are dealing with an inherently nonstationary process that cannot be represented as a Poisson process with rates proportional to time (*cf.*, see (1.6) below) (Lomnitz and Hax, 1966).

3) The magnitude of the largest aftershock, m_n, is given in terms of the magnitude of the main shock, $m^{\star}$, according to Båth's law

$$m_n = m^{\star} - 1.2. \qquad (1.5)$$

Båth's law is an explicit statement that there is an upper limit to the magnitude of an aftershock. The magnitude of the largest aftershock is related to it in a definite manner that neither depends on the magnitude of the main shock nor, in any way, depends on the particular nature of the aftershock sequence.

Time

1) The frequency of aftershocks decays in time according to the modified Omori relation (Utsu, 1969-1970)

$$\dot{N}(t) = At^{-p} \qquad (1.6)$$

where t is time (days), A is a numerical parameter, and $p \geq 1$. It coincides with Omori's law for $p = 1$ (Omori, 1984; Utsu *et al.*, 1995).

2) The rate of energy release per unit time also follows a hyperbolic law

$$\dot{\varepsilon}(t) = Bt^{-q} \qquad (1.7)$$

where B is a numerical parameter, and $q > p$. The rate of energy decrease is more rapid than the decrease in the frequency of aftershocks (Utsu, 1969-1970).

Volume

1) On the basis of the assumption that the energy which generates the aftershocks remains in the source volume of the main shock, there is a linear relation between this volume $V^{\star}$ (cm^3) and the energy of the main shock (Utsu, 1969-1970), namely

$$\varepsilon^{\star} = \eta^{\star} V^{\star} \qquad (1.8)$$

where $\eta^{\star}$ is the constant energy density. Based on empirical evidence, Båth and Duda (1964) proposed that the volume affected by an aftershock, V, is proportional to the energy released, $\varepsilon = \eta V$, where η is a constant, pertaining to the aftershock energy density. In other words, $V/V^{\star}$ should have the same distribution as $\varepsilon/\varepsilon^{\star}$.

If the asymptotic magnitude distribution of aftershocks has a right endpoint, their initial distribution will be given by the one-parameter beta distribution

$$F(\varepsilon) = \left(\frac{\varepsilon}{\varepsilon^{\star}}\right)^{\tilde{\rho}}. \qquad (1.9)$$

The exponent $\tilde{\rho}$ is called the «clustering dimension» (Hastings and Sugihara, 1993). A power law scaling is used to define an object consisting of a finite number of points. The object is to be covered by boxes of size ℓ which is much smaller than the side of the box enclosing all the points, L. If we are considering an area in which the points are located, the large box can be divided up into L^2/ℓ^2 small boxes of size ℓ. The scaling range in this case is 2. Likewise, if we consider points along a line of length L, then the interval can be covered by L/ℓ smaller boxes with a box dimension equal to 1. In general, the average number of points in a box of size ℓ scales as ℓ^D, where the exponent D is the cluster dimension.

If aftershocks do constitute a compound Poisson process (Vere-Jones, 1966), (1.9) can be thought of as replacing it by a more general, continuous, «pure jump» process with stationary increments. Moreover, if we regard the aftershock sequence as a set of *order-statistics*, consisting of independent and identically distributed random variables that are arranged in order of energy, the members of the ordered set are no longer independent, nor identically distributed (Stuart and Ord, 1994). However, if the unordered sample has a distribution function, $F(x)$, with a continuous density $f(x)$, the new distribution of the ordered set will be expressible in terms of the original distribution according to a two-parameter beta distribution (Stuart and Ord, 1994). Moreover, if the size of the sequence is sufficiently large, we should expect the upper order-statistics to be independent of the lower order-statistics, and both will tend to independent gamma distributions (Cramér, 1946). That is to say, as the number of aftershocks increases, the joint probability distribution for the largest and smallest energies will reduce to a product of independent distributions.

The two-parameter beta distribution was used by Utsu (1969-1970) to calculate the average energy of the rth largest earthquake in terms of the energy of the largest aftershock. This implies that F must have a right endpoint which is incompatible with the Pareto distribution, which has a left endpoint but no right endpoint. And since the GR law is related to the Pareto law by the magnitude-energy relation (1.4), it puts into doubt the validity of the GR law in providing a complete characterization of the entire aftershock sequence. Hence, it is still an open question what is the correct sampling distribution to use, and whether this distribution will be the same for both upper and lower order-statistics of the aftershock sequence.

2. Order-statistics and aftershock sequences

In this section we will show that in certain aftershock sequences the distributions regarding the top and the bottom order-statistics may be different. The idea that aftershocks should follow order-statistics is implicit in the work of Utsu (1969-1970), who used the two-parameter beta distribution to calculate the average energy of the rth largest aftershock. Assuming the aftershocks were infinite in number, Utsu found that the total energy released by the aftershocks is proportional to the energy of the largest aftershock. Specifically, Utsu set $(\varepsilon_r/\varepsilon_n)^{\tilde{\beta}} = 1/r$, where ε_n is the energy of the largest aftershock, and expressed the sum

$$\mathcal{E} = \sum_{r=1}^{\infty} \varepsilon_r = \varepsilon_n \sum_{r=1}^{\infty} r^{-1/\tilde{\beta}} = \varepsilon_n \zeta(\beta/\tilde{b})$$

in terms of the Riemann ζ-function. For $\beta = \frac{3}{2}$ and $\tilde{b} = 1$, $\zeta(\frac{3}{2}) = 2.6$ so that the total energy released, $\mathcal{E}$, in the aftershock sequence is 2.6 times the energy of the largest aftershock, ε_n. However, Utsu realized that if fluctuations were taken into account, the average energy of the rth aftershock should be determined from a two-parameter beta distribution, with an *a priori* probability given by (1.9) [1]. The upper limit on the aftershock energy, according to Utsu, is expressible in terms of the energy of the largest aftershock, and not the main shock.

Consider an aftershock sequence of n distinct values which are arranged in ascending order so that we can talk about the rth largest

[1] The one-parameter beta distribution (1.9) can be considered as a special case of the two-parameter beta distribution

$$\frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}$$

namely for $b = 1$, where $B(a,b)$ is the beta function (Reiss and Thomas, 1997).

value from the bottom (2). If the number of aftershocks in the sequence is sufficiently large so that there is a statistical independence between aftershocks at the top and bottom of the sequence, we will show that the aftershocks at the top will be governed by the one-parameter beta distribution, (1.9), while those at the bottom by the Pareto distribution, (1.1).

Considering aftershocks at the top, let $F(m) = \boldsymbol{Nr}(\leq m)/n$, where $\boldsymbol{Nr}(\leq m)$ is the number of aftershocks having a magnitude m or less. If the one-parameter beta distribution (1.9) applies to the distribution in energy then the regression law, in terms of the magnitude, will be given by Lavenda and Cipollone (2000)

$$\log \boldsymbol{Nr}(\leq m) = \tilde{a} + \tilde{b}m \tag{2.1}$$

where

$$\tilde{a} = \log n - \tilde{b}m^\star \tag{2.2}$$

and $m^\star$ is the magnitude of the main shock. We will refer to (2.1) as the beta (B) regression law, and compare it to the GR law (1.2).

Let E_r denote the rth value of the energy from the bottom. The probability $g_r(\varepsilon)\,d\varepsilon$ that $n-r$ energies lie above ε, and $r-1$ lie below it, while the remaining value falls between ε and $\varepsilon + d\varepsilon$ is given by the two-parameter beta distribution

$$g_r(\varepsilon)d\varepsilon = \frac{F^{n-r}(\varepsilon)(1-F(\varepsilon))^{r-1}f(\varepsilon)d\varepsilon}{B(r,n-r+1)} \tag{2.3}$$

$$= n\binom{n-1}{r-1}\left(\frac{x}{n}\right)^{n-r}\left(1-\frac{x}{n}\right)^{r-1} d\frac{x}{n} = g_r(y)\,dy$$

where $F(\varepsilon)$ is the one-parameter beta distribution, (1.9), f is the density of F, and B is the beta function. Utsu (1969-1970) considered the distribution (2.3) for $r = n-1$, because he set the energy of the largest aftershock as the upper limit on the energy. In this case, the two-parameter beta distribution (2.3) reduces to

$$g_{n-1}(y)\,dy = n(n-1)y(1-y)^{n-2}\,dy$$

where $y = (\varepsilon/\varepsilon^\star)^{\tilde{\rho}}$ and $\varepsilon^\star$ is now to be regarded as the energy of the largest aftershock. Using this probability distribution, Utsu determined the average energy of the nth aftershock as

$$\boldsymbol{E}(E_r) = \varepsilon^\star n(n-1)\int_0^1 y^{1+1/\tilde{\rho}}(1-y)^{n-2}\,dy =$$

$$= \varepsilon^\star \frac{1+\tilde{\rho}}{\tilde{\rho}^2}B(n+1, 1/\tilde{\rho}).$$

The total energy,

$$\mathcal{E} = \sum_{n=0}^{\infty}\boldsymbol{E}(E_n) = \left(\frac{1+\tilde{\rho}}{1-\tilde{\rho}}\right)\frac{\varepsilon^\star}{\tilde{\rho}}$$

is then obtained by summing over all n, and noting that $\sum_{n=0}^{\infty}B(n+1, 1/\tilde{\rho}) = B(1, 1/\tilde{\rho}-1)$.

There are two important points concerning Utsu's derivation. Firstly, in order to use the one-parameter beta distribution (1.9) in the two-parameter beta distribution (2.3), Utsu had to switch the distribution for its tail. That is, the probability of finding $n-r$ values above ε was taken as y^{n-r}, and not $(1-y)^{n-r}$ (Cramer, 1946). Although the one-parameter beta distribution (1.9) is the initial distribution for the smallest value (Reiss and Thomas, 1997), it is precisely this distribution that enabled Utsu to express the total energy released in an (infinite) aftershock sequence in terms of the energy of the largest aftershock. Had the tail Pareto distribution (1.1) been used, it would have expressed the total energy in terms of the minimum detectable energy in the aftershock sequence (*cf.* (2.16) below). Secondly, the inequality $\tilde{\rho} < 1$ will be seen in Section 5 to be precisely the condition for the existence of a concave entropy function.

Since we are considering the upper limit of the energy as the energy of the main shock, $\varepsilon^\star$, the distribution in the energy of the largest aftershock is obtained by setting $r = n$. Then, in the limit of a large sequence, the two-parameter

(2) If there is more than one aftershock with the same magnitude then any one of them can be considered as the order-statistic of that magnitude.

beta distribution (2.3) becomes the exponential distribution

$$g_n(x)dx = e^{-x}dx. \qquad (2.4)$$

And because the initial distribution is the one-parameter beta distribution, (1.9), the exponential distribution (2.4) is actually the Weibull distribution:

$$g_n(\varepsilon)d\varepsilon = n\tilde{\rho}\frac{\varepsilon^{\tilde{\rho}-1}}{\varepsilon^{\star\tilde{\rho}}}e^{-n(\varepsilon/\varepsilon^\star)^{\tilde{\rho}}}d\varepsilon. \qquad (2.5)$$

The Weibull distribution predicts that the mode and mean of the energy of the largest aftershock will decrease as $n^{-1/\tilde{\rho}}$, while the variance will decrease as $n^{-2/\tilde{\rho}}$.

The Weibull energy distribution is converted into the Gompertz magnitude distribution:

$$g_1(m)dm = n\tilde{b}'e^{\tilde{b}'(m-m^\star)}\exp(-ne^{\tilde{b}'(m-m^\star)})dm \qquad (2.6)$$

by the energy-magnitude relation (1.4), where $\tilde{b}' = \tilde{b}/\log e$. It is the only distribution that has an exponential hazard rate

$$h(m) = \frac{f(m)}{1-F(m)} = n\tilde{b}'e^{\tilde{b}'(m-m^\star)}. \qquad (2.7)$$

The Gompertz distribution (2.6) is the distribution of ages at death, and (2.7) gives the mortality rate. The mean magnitude of the largest earthquake

$$\boldsymbol{E}(M_n) = m^\star - \tilde{b}'^{-1}(\gamma + \ln n) \qquad (2.8)$$

bears a remarkable resemblance to Båth's law (1.5), where $\gamma \approx 0.577$ is Euler's constant. The difference between (2.8) and Båth's law (1.5) is the dependence of the mean largest aftershock on the logarithm of the number of aftershocks in the sequence. According to Båth's law, the difference in the instrumental magnitudes of the main shock and the largest aftershock in a sequence is 1.2, independent of both the magnitude of the main shock and the nature of the sequence. Assuming that the magnitudes in an aftershock sequence are distributed according to an exponential distribution, Vere-Jones (1969) found that the number of aftershocks should vary as the exponential of the magnitude of the main shock, or, equivalently, that expected magnitude of the main shock should vary as the logarithm of the number of aftershocks. Also there is some observational evidence that the number of aftershocks varies as the square root of the energy of the main shock (Solov'ev and Solov'eva, 1962).

The probability distribution (1.9) gives the average energy of the rth largest aftershock in terms of the energy of the main shock as

$$\boldsymbol{E}(E_r) =$$

$$= n\varepsilon^\star\binom{n-1}{r-1}\int_0^1\left(\frac{x}{n}\right)^{r-1+1/\tilde{\rho}}\left(1-\frac{x}{n}\right)^{n-r}d\frac{x}{n}$$

$$= \varepsilon^\star\frac{\Gamma(r+1/\tilde{\rho})}{\Gamma(r)}\frac{\Gamma(n+1)}{\Gamma(n+1+1/\tilde{\rho})}. \qquad (2.9)$$

For the Aleutian Islands 1957 aftershock sequence, to be discussed in the next section, $\tilde{\rho} \approx 0.5$. The average energy of the rth quake from the bottom (2.9) becomes

$$\boldsymbol{E}(Er) = \frac{r(r+1)}{(n+1)(n+2)}\varepsilon^\star. \qquad (2.10)$$

The total energy $\mathcal{E}$ released in the aftershock sequence is the sum of (2.10), namely,

$$\mathcal{E} = \sum_{r=1}^{n}\boldsymbol{E}(E_r) = \frac{1}{3}n\varepsilon^\star$$

giving an average aftershock energy of $\varepsilon^\star/3$. Since $m^\star = 8.3$, the average magnitude would be 7.98 which is approximately 1.3 times greater than the calculated value, 6.19, obtained by partial averaging of groups of aftershocks (Ranalli, 1969). Rather, if the energy of the largest aftershock is used, the average magnitude becomes 6.98, which is still 1.1 times larger than the calculated value. The reason for this discrepancy is that the distribution used to determine the expected energy value applies to order-statistics for large values of r, and not to small values.

The Pareto tail distribution (1.1) can be thought of as the intensity of small jumps that give rise to a compound Poisson distribution (de Finetti, 1970). In this interpretation, the exponent ρ must be less than 2 in order that the variance of the small jumps be finite. If $\rho < 1$ only positive jumps occur, whereas for $\rho > 1$ both positive and negative jumps take place. According to the usual quoted value of $\rho = b/\beta = \frac{2}{3}$ (Utsu, 1969-1970), it would appear that only small *positive* jumps occur. The process of energy decay entails a jump in the energy to a lower value after each aftershock (Vere-Jones, 1966). Each aftershock represents a jump in the process, where both the size of the jumps and their intensity are regulated by the Pareto law (1.1).

We now turn our attention to the order-statistics at the bottom. In the usual treatment, the orderings are made from the top and the bottom separately (Cramér, 1946). However, we still keep the same ordering: $r = n$ corresponds to the largest order-statistic and $r = 1$, to the smallest. The probability that $n - r$ values lie above ε, having an *a priori* probability of $1 - y$, $r - 1$ values lie below ε, with probability y, and the remaining value lying between ε and $\varepsilon + d\varepsilon$ is

$$g_r(y)dy = n\binom{n-1}{r-1}y^{r-1}(1-y)^{n-r}dy \qquad (2.11)$$

where $y = (\varepsilon_\star/\varepsilon)^\rho$, and $\varepsilon_\star$ is the minimum detectable energy of the aftershock sequence. In the limit of a large sequence, (2.11) becomes the gamma distribution

$$g_r(x)dx = \frac{x^{r-1}}{\Gamma(r)}e^{-x}dx \qquad (2.12)$$

where $x = ny$, for the rth order-statistic from the bottom. Unlike the asymptotic exponential distribution (2.4) of the two-parameter beta distribution (2.3), which is valid for the largest aftershock, (2.12) is an order-statistic distribution for $r > 1$. This shows that the GR regression law (1.2) should be valid for a greater number of smaller aftershocks and the B law for a lesser number of large ftershocks. In other words, there will be many more smaller shocks that will be attracted to the asymptotic distribution (2.12) than larger shocks to the symptotic form of (2.3). Strictly speaking, the only asymptotic distribution that coincides with an actual distribution is the exponential distribution (2.4), which is the distribution of the energy of the largest aftershock.

Converting the gamma distribution (2.12) into an energy distribution

$$g_1(\varepsilon)d\varepsilon = n\rho\frac{\varepsilon_\star^\rho}{\varepsilon^{\rho+1}}e^{-n(\varepsilon_\star/\varepsilon)^\rho}d\varepsilon \qquad (2.13)$$

which is known as the Fréchet distribution. According to (2.13), the most probable value of energy *increases* as $n^{1/\rho}$ with the number of aftershocks in any given sequence. This is to be compared with the most probable value of the energy of the Weibull distribution (2.5) which decreases as $n^{-1/\beta}$. Transforming from energy to magnitude, the Fréchet distribution (2.13) becomes the double exponential, or Gumbel, distribution

$$g_1(m)\,dm = nb'e^{-b'(m-m_\star)-ne^{-b'(m-m_\star)}}dm \qquad (2.14)$$

which has a mean value of

$$\boldsymbol{E}(M_1) = m_\star + b'^{-1}(\gamma + \ln n) \qquad (2.15)$$

A comparison of (2.8) and (2.15) will serve to justify our choice of the *a priori* probabilities in the two-parameter beta distributions, (2.3) and (2.11). The location parameters in the former and latter distributions are $m^\star$ and $m_\star$, respectively. Since they physically represent the magnitude of the main shock and the minimum detectable magnitude in the sequence, it would hardly make sense to have a mean magnitude greater than the magnitude of the main shock, or a magnitude less than the minimum detectable magnitude. And because the Gompertz and Gumbel distributions are 'mirror' images of one another (*i.e.*, under the formal transformation $m \to -m$), (2.8) predicts that the mean value of the largest aftershock will decrease as the logarithm of the number of aftershocks of different magnitude, while (2.15) predicts that the mean value of the smallest aftershock will increase by the same amount. Their affect on the mean values is regulated by the size of the scale parameters, $\tilde{b}'^{-1}$ and b'^{-1}, in the two distributions.

Moreover, the Gumbel distribution (2.14) can be thought of as a Poisson process

$$\Pr\left(\max_{1\le i\le n} M_i \le m\right) = e^{-n\lambda_m}$$

with an exponential rate $\lambda_m = e^{-b'(m-m_\star)}$ (Aldous, 1989). This explains why the GB law works as well as it does: for large m, the maximum M behaves as if all the magnitudes of the aftershocks were independent and identically distributed. That is, they are insensitive to the magnitude of the main shock, having values which are appreciably smaller than the main shock.

Another reason why the Fréchet distribution (2.13) will not provide a global description of the energetics of an aftershock sequence is that it gives the average energy for the rth order-statistic from the bottom in terms of the minimum detectable energy, namely

$$E(E_r) = \varepsilon_\star \frac{\Gamma(n+1)}{\Gamma(n+1-1/\rho)} \frac{\Gamma(r-1/\rho)}{\Gamma(r)}. \quad (2.16)$$

If we attempt to define the total energy release as the sum of (2.16) we would be immediately confronted by the fact that it would be given in terms of the minimum detectable energy of an aftershock. This is hardly a satisfactory account of the total energy release since there is no information on the energy of the main shock. This is yet a further limitation on validity of the GR law in the analysis of aftershock sequences.

The above conclusions can be formulated in the testing of the hypothesis of whether the aftershock sequence consists of a set of independent and identically distributed random variables against the alternative hypothesis of whether they exhibit clustering. The generalized Omori law (1.6) and the rate of energy release (1.7) can be considered as expressing the dependence of each aftershock on the main shock (Utsu, 1969-1970; Lomnitz and Hax, 1966). Clustering would imply that the dependence between aftershocks is significant in terms of the energy of the main shock, while their independence would imply that the aftershocks are so weak that the influence between two successive shocks on the energy of the main shock is negligible. In the next section we analyze the Aleutian Islands 1957 aftershock sequence, where large magnitude aftershocks do show a dependency on the magnitude of the main shock. Other analyses, like the 1957 San Francisco aftershock sequence showed negligible dependence of on the magnitude of the main shock. This is in agreement with the findings of Lomnitz and Hax (1966). In the San Francisco aftershock sequence the GR law was found to apply to the entire sequence.

3. Analysis of an aftershock sequence

The reason why the GR law seems to work so well is that aftershocks of small magnitude have a preponderate effect on the sequence. Once these aftershocks have been excluded from the sample, a different picture may emerge: large energy aftershocks may appear to follow the one-parameter beta law (1.9), indicating that the appropriate regression law is (2.1) and not (1.2) for such high energy aftershocks. We base our statistical analysis on:

Regression analysis – The R^2 statistic, or the coefficient of determination, is a measure of the dependent's variability that is explained by the independent variable (Abacus Concepts, 1994). The closer R^2 is to 1, the better the accountability. Since the number of independent variables is not included in the R^2 statistic, its value is bound to increase as more independent variables are added. The «adjusted» R^2 statistic is used to remedy this situation.

Mean excess function – The mean excess function is the mean of the magnitudes greater than some prescribed threshold value m_0. In other words, it is the first moment of the distribution of m given that $m \ge m_0$. Only for the generalized Pareto distribution is the mean excess, or mean residual life, function a straight line. It is a linear increasing function for the Pareto distribution (1.1) with shape parameters $\rho > 1$. In contrast, it is a linearly decreasing function for the one-parameter beta distribution (1.9) (Reiss and Thomas, 1997).

Hazard function – The hazard function, or the mortality rate, is the derivative of the residual life distribution at any given age. It is a decreasing function for the Pareto distribution (1.1) for shape parameters $\rho > 1$, whereas it is an increasing function for the one-parameter beta distribution, (1.9) with shape parameters $\tilde{\rho} > 1$.

We now consider a case of an aftershock sequence taken from the catalogue in Ranalli (1969) in which the GR law does not provide an adequate regression law for large magnitude aftershocks.

Aleutian Islands 1957

The main shock had a magnitude $m^\star = 8.3$. The minimum detectable magnitude was $m_\star = 5.85$. The number of aftershocks recorded in the first 100 days was 205, with $n = 15$ distinct magnitudes ranging from 5.9 to 7.3.

– Figure 1: *regression analysis* (Abacus Concepts, 1994). The GR law (1.2) (left column) is compared with the regression law of the B law (2.1) (right column). The upper left-hand plot has no events excluded, and the GR law has a $R^2 = 0.962$ and an adjusted $R^2 = 0.959$. The cal-

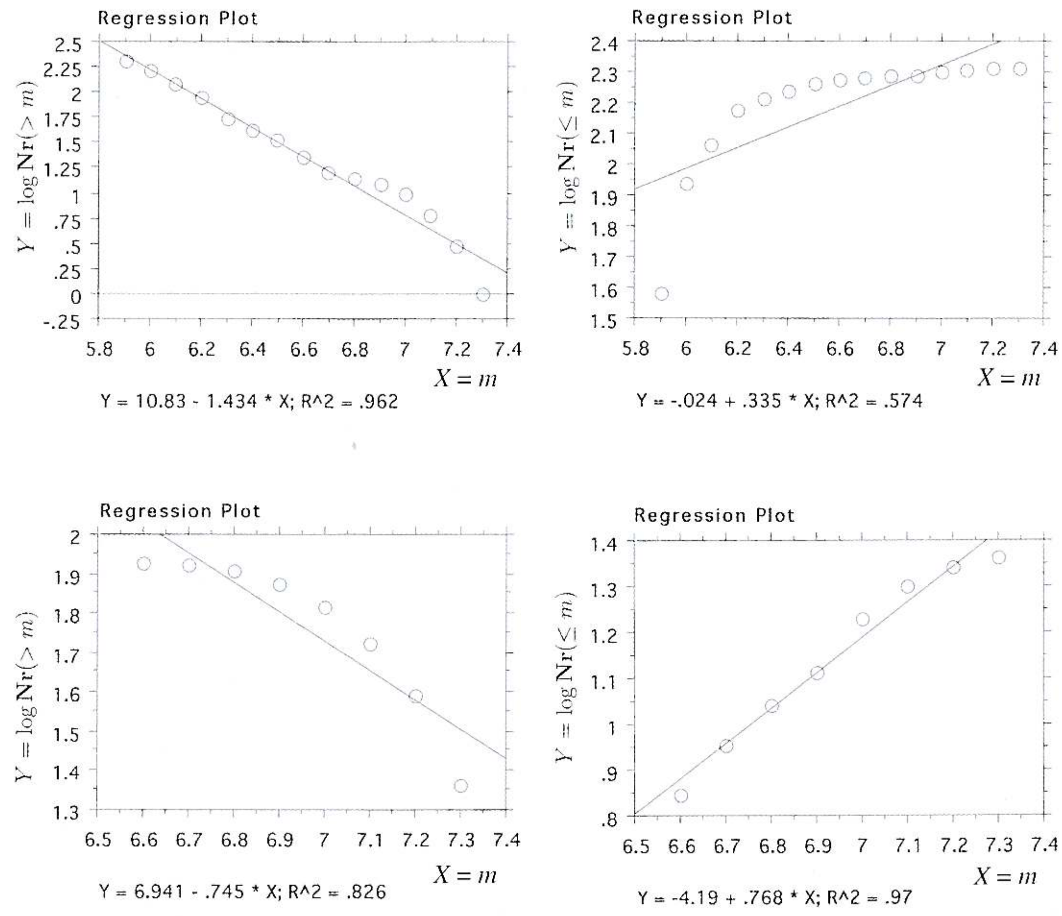

Fig. 1. The GR regression law shown in the left column is compared to the B regression law in the right column without (first row) and with (second row) the exclusion of shocks with magnitudes less than 6.5.

culated value of the intercept, determined from (1.3), is $a = 10.77$ which is to within .5% of the value 10.83. The B law, shown on the upper right-hand side, has a $R^2 = 0.574$ and an adjusted $R^2 = 0.542$. There is no comparison between the calculated intercept value and the value of the linear regression curve. On the bottom left-hand side, aftershocks whose magnitude were less than $m = 6.5$ were excluded. The GR law now has a $R^2 = 0.826$, while on the lower right-hand side, the B law has a $R^2 = 0.97$. The intercept, determined from (2.2) is $\tilde{a} = -4.06$ which is to within 3% of the intercept value, -4.19.

– Figure 2: in the top figure, the *sample mean excess function* shows two diametrically opposing tendencies, separated by an energy threshold which corresponds to a magnitude of 6.5. In the bottom figure, the *sample hazard function* is plotted as a function of the energy. For the full sample data the sample hazard function decreases, while the exclusion of aftershocks whose energies correspond to magnitudes smaller than $m = 6.5$ actually inverts its tendency and the sample hazard function increases with energy. This indicates that there is a long-living subpopulation that becomes visible beyond the magnitude of 6.5.

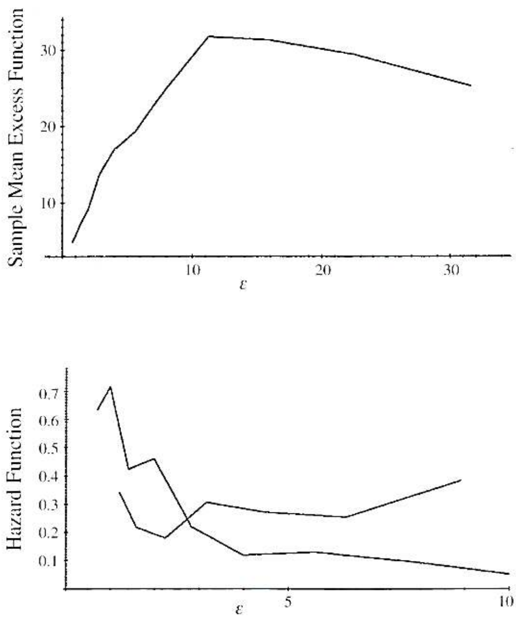

Fig. 2. The sample mean excess function in the top figure is plotted against the energy. The hazard function shown in the bottom figure is also plotted as a function of the energy. The increasing curve excludes those energies corresponding to magnitudes less than 6.5.

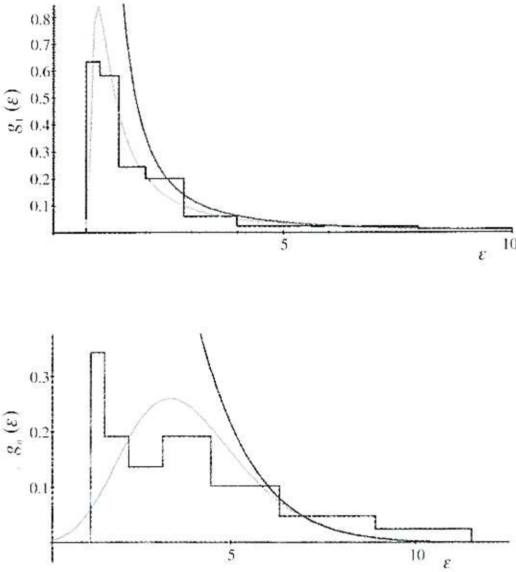

Fig. 3. In the top figure, the Fréchet density is fitted to the full histogram, while excluding shocks whose energies correspond to magnitudes less than 6.5 leads to the fitting of a Weibull density with a right endpoint to the histogram curve in the bottom of the figure. The Pareto tail distribution is shown in both figures for comparison.

– Figure 3: the top diagram is the *histogram* of the entire aftershock sequence. A Fréchet distribution, with positive shape parameter in the terminology of Reiss and Thomas (1997), is fitted to the histogram. The Pareto tail distribution is shown for comparison. The histogram in the bottom diagram excludes aftershocks corresponding to magnitudes less than 6.5. The shape parameter, in the sense of Reiss and Thomas (1997), is negative enabling a Weibull density with a right endpoint to be fitted to the histogram. Again the Pareto tail distribution is shown for comparison. The plots were obtained using the statistical software XTREMES (Reiss and Thomas, 1997).

– Figure 4: the Fréchet and Pareto quantile functions (smooth curves) are fitted to the sample quantile function (jagged curve) in the top figure using the complete set of data. The quantile functions for the Pareto and Fréchet distributions are overlapping, having an exponent of nearly one, location parameter almost zero, and a scale parameter of 3. The sample quantile function shows the same convex behavior as the Fréchet quantile function. In the bottom figure, the quantile function of the Weibull distribution (smooth curve), with an exponent 3, and location and scale parameters 10 and 7, respectively, is fitted to the sample quantile function of the excluded data set. The fit is quite good.

Aftershocks with magnitudes greater than 6.5 in the Aleutian Islands aftershock sequence of 1957 manifest a dependency on the magnitude of the main shock and hence tend to cluster, whereas aftershocks below 6.5 do not, and

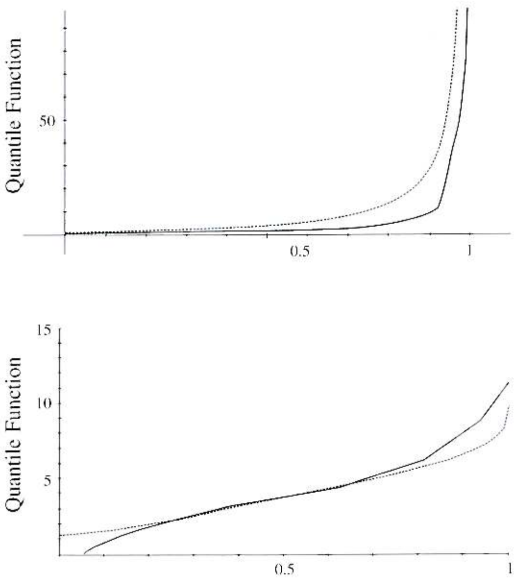

Fig. 4. In the top figure, the sample quantile function (jagged curve) is fitted to the overlapping Fréchet and Pareto quantile functions (smooth curves) using all the data from the Aleutian aftershock sequence. In the bottom figure, events of magnitude of less than 6.5 are excluded and the quantile function of a Weibull distribution (smooth curve) is fitted to the sample quantile function (jagged curve).

consequently are independent. Other examples, such as the San Francisco 1957 aftershock sequence, showed no clustering at all. This has been confirmed elsewhere (Lomnitz and Hax, 1966). Aftershocks were insensitive to the main shock possibly due to the shallowness of the main shock. The GR law gave a good regression fit over the entire sequence.

4. Rates and returns

4.1. *Nonstationary and inhomogeneous point processes*

Aftershocks are notoriously nonstationary processes because their rates are an explicit function of time. They can neither pretend to be a set of independent and identical set of random variables, discrete (M_n; $n > 1$) or continuous (M_t; $t > 0$), so that their maximum values need not coincide with $M_n = \max_{1 \le i \le n} M_i$, or $M_t = \sup_{0 \le s \le t} M_s$.

A relation between these two descriptions of aftershock sequences can be obtained by defining what is known as a «hitting time»

$$\boldsymbol{T}_m = \min \{t : M_t > m\}$$

or the first time that the process exceeds a given magnitude m. Consequently, the probability that the magnitude m will not be exceeded in time t is equivalent to the probability that the hitting time will exceed t

$$\Pr (M_t \le m) = \Pr (\boldsymbol{T}_m > t). \qquad (4.1)$$

Now, if m and t are sufficiently large, the probability that an aftershock is greater than m given that a preceding one was greater than m tends to zero. Hence, the maximum aftershock M_n is asymptotically the same as that of an independent and identically distributed set of random variables (Aldous, 1989).

A stationary process will have a probability of survival that decays exponentially with time

$$\Pr (\boldsymbol{T}_N > t) \approx \exp(-\lambda_N t) \qquad (4.2)$$

as in the case of a Poisson process, where the rate λ_N is constant. Thus, the waiting time has approximately an exponential distribution with a mean waiting time $1/\lambda_N$. However, fig. 5 shows that aftershocks in the Aleutian Islands sequence follow Omori's law (1.6) with an exponent $p \approx 1$. The actual value $p = 0.997$ agrees exactly with the least square estimate (Ranalli, 1969). We thus have a nonstationary process so that (4.2) must be replaced by

$$\Pr(T_N > t) \approx \exp\left(-\int_1^t \dot{N}(s)ds\right) = t^{-A} = e^{-N(t)}$$

where $N(t)$ is the cumulative number of aftershocks up to time t. The aftershocks occurring during the first day have been excluded because of their possible incompleteness with respect to the number of shocks counted, due to the high frequency of aftershocks (Ranalli, 1969).

For non-homogeneous «space-time» point processes, the probability of the magnitude of the largest aftershock M_n being less than some fixed value m can be expressed in terms of the initial distribution of the largest value (Aldous, 1989). If ϕ is the intensity function, such that $\phi(m)dm\ dt$ is the chance of a point falling in $[m, m + dm] \times [t, t + dt]$, then

$$(4.1) \approx \exp\left(-t\int_m^\infty \phi(x)dx\right) \tag{4.3}$$

is the probability of there being no points in the interval (m, ∞) in the time interval $(0, t)$. The intensity function ϕ is the derivative of the exponential tail density, $b'e^{-b'(x-m_\star)}$. Equating the rates of the non-homogeneous and nonstationary process in (4.3) results in Omori's law rewritten in probability terms

$$\dot{N}(t) = \Pr(> m) \tag{4.4}$$

where $\Pr(> m)$ is the initial exponential tail distribution. In the case where clustering occurs, the right-hand side of (4.4) has to be multiplied by the probability that, in the short term, there is only one magnitude exceeding m. This has been referred to as the «extremal index» of the process (Leadbetter and Rootzen, 1988). This is yet another indication that the GR law has built into it the property that the aftershocks behave as a set of independent and identically distributed

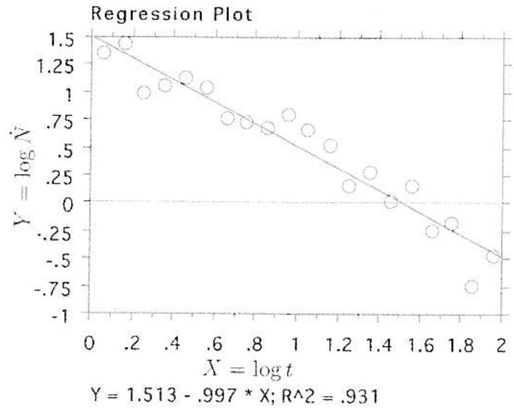

Fig. 5. Omori's regression law.

random variables. Expression (4.4) is explicitly given as

$$m - m_\star = b^{-1} \log(t/A)$$

which is comparable to (2.15) with time interval t replacing the size of the aftershock sequence, n. Furthermore, it is entirely compatible with the GR law; introducing GR law on the left-hand side gives $\Pr(> m) = A/t$, which is (4.4).

Now let $M_t = \min_{0 \le s \le t} M_s$ be the minimum of a set of independent and identically distributed random variables. Instead of (4.3) we now have

$$(4.1) \approx 1 - \exp\left(-t\int_0^m \tilde{\phi}(x)dx\right)$$

for the probability of a point falling inside the interval $(0, m)$ in the time interval $(0, t)$. The process has intensity, $\tilde{\phi} = \tilde{b}'e^{\tilde{b}'(x-m^\star)}$, and rate

$$\dot{N}(t) = \Pr(\le m). \tag{4.5}$$

With the rate given by Omori's law, (4.5) becomes

$$m^\star - m = \tilde{b}^{-1} \log(t/A)$$

which is comparable to (2.8) with the time interval replacing the number of aftershocks in the sequence. The former asserts that the difference between the magnitude of the main shock and the mean of the largest aftershock increases as the logarithm of the number of aftershocks in the sequence. The latter says that the difference increases as the logarithm of the time elapsed since the main shock occurred. Introducing the B law (2.1) into the left-hand side gives back the rate (4.5).

4.2. *Return periods*

The mean return period, or the mean interval in days, between aftershocks having a magnitude greater than m is

$$\boldsymbol{E}(\boldsymbol{T}_m)=\int_0^\infty t\,d\,\Pr(\boldsymbol{T}_m\le t)=n/\boldsymbol{Nr}(>m)= \tag{4.6}$$

$$=\begin{cases}\exp[b'(m-m_\star)]\\ [1-e^{-\tilde{b}'(m^\star-m)}]^{-1}\end{cases}$$

according to the Gumbel (2.14), and Gompertz (2.6) distributions, respectively. In words, (4.6) states that if an event has probability p of occurring, it will take, on the average, $1/p$ trials before that event will happen once (Gumbel, 1958). Both the GR and B laws confirm that the mean return time of aftershocks occur like main shocks: the greater the magnitude the longer is the mean return period. However, the B law attributes a much shorter return period for larger aftershocks and a larger return period for aftershocks of smaller magnitude, as can be seen from table I of return periods from the Aleutian Islands aftershock sequence. The slope of the GB and B laws were taken to be $b = 1.434$ and $\tilde{b} = 0.768$, respectively, as calculated from the non-excluded and excluded data sets, respectively. The value of $b = 1.434$ is to within 12% of the maximum likelihood estimate $b = 1.277$ (Ranalli, 1969). The return period for the B law appears quite sensitive to whether the magnitude of the main shock or the magnitude of the largest aftershock is used. Opting for the latter has the effect of increasing the return period of both smaller and larger aftershocks considerably, as can be seen from a comparison of the last two columns in table I. Since the increase in the return period of smaller aftershocks does not conform to observation, the return period should be calculated using the magnitude of the main shock.

Table I. Return periods calculated from the Aleutian aftershock sequence.

m	$\boldsymbol{Nr}(\ge m)$	$\boldsymbol{T}_m$(GB)	$\boldsymbol{T}_{m^\star}$(B)	$\boldsymbol{T}_{m_n}$ (B)
7	10	35.03	6.62	26.96
6.6	23	9.61	5.19	11.85
6.5	33	6.96	4.93	10.42
6.3	55	3.64	4.49	8.45
6.2	89	2.64	4.13	7.73

Stability in time is reflected in small oscillations about the mean magnitude. In order to eliminate large fluctuations, the mean magnitude of each group of 10 aftershocks was calculated by Ranalli (1969). In the case of the Aleutian Islands aftershock sequence, the oscillations occurred about a mean magnitude of $\boldsymbol{E}(M) = 6.19$, and there was no sign of decay in time for a period of 100 days. The number of partial mean magnitudes was 29, and 100% of the oscillations occurred within $\boldsymbol{E}(M) \pm 0.20$. The reason why the GR average return period seems to be closer to the calculated value is that the partial averaging, by grouping of 10 successive aftershocks, tends to decrease the importance of large values of the magnitudes of the aftershocks, especially in the first few days after the main shock.

5. Thermodynamics of aftershocks

Like the particle number of degenerate gases, the number of aftershocks is not a conserved quantity, but, rather, varies with energy or temperature. A photon gas, obeying the statistics of black-body radiation, is a good example of a degenerate gas. And like a photon gas, the entropy of an aftershock sequence will be proportional to the number of aftershocks (Lavenda, 1991)

$$S=\frac{n}{\tilde{\rho}}\left(\frac{\varepsilon}{\varepsilon^\star}\right)^{\tilde{\rho}}=\tilde{\rho}^{-1}N(t) \tag{5.1}$$

where $N(t)$ is the integral of Omori's law (1.6). In order that the entropy be a concave function of the energy, $\tilde{\rho} < 1$ (Lavenda, 1995), which confirms Utsu's result. In comparison to black-body radiation, where $\tilde{\rho} = 0.75$, the Aleutian Islands earthquake sequence has an energy exponent of $\tilde{\rho} = \tilde{b}/\beta = 0.768/1.5 = 0.512$.

The entropy production produced by the decaying process of aftershocks is

$$\dot{S} = \frac{n}{\varepsilon^\star}\left(\frac{\varepsilon}{\varepsilon^\star}\right)^{\tilde{\rho}-1}\dot{\varepsilon} = \frac{A}{\tilde{\rho}t} \tag{5.2}$$

where Omori's law (1.6), with $p = 1$, has been used, since it applies to the Aleutian Island aftershock sequence. Whereas the entropy production decays hyperbolically in time (Lomnitz, 1994), the rate of energy release of an aftershock decays as

$$\tilde{\rho}\frac{d}{dt}\ln\varepsilon = \frac{1}{t\ln t}. \tag{5.3}$$

Since $\tilde{\rho} < 1$, the energy released increases as $(\ln t)^{1/\tilde{\rho}}$, which is faster than that predicted by Omori's law, which increases as $\ln t$ (Lomnitz, 1994). Hence, the rate of energy release in an aftershock will be given by

$$\dot{\varepsilon} \sim (\ln t)^{(\tilde{\rho}^{-1}-1)}/t \tag{5.4}$$

in contrast to Utsu's relation (1.7) with exponent of $q \approx 2$. The rate of energy release in the Aleutian Islands aftershock sequence, $\ln t/t$, would tend to zero not nearly so rapidly as Utsu's law (1.7) predicts. All powers of $\ln t$ vary slowly at $t = 0$ and $t = \infty$ meaning that scaling does not affect the function at these extremes.

According to the second law of thermodynamics, the inverse temperature is defined as $T^{-1} = \partial S/\partial\varepsilon$, and this gives the thermal equation of state

$$\varepsilon(T) = (\mathrm{nT})^{1/(1-\tilde{\rho})}/\varepsilon^{\star\,\tilde{\rho}(1-\tilde{\rho})}. \tag{5.5}$$

Kinetic theory equates the temperature with the average kinetic energy with which the molecules in a system move about. The fact that the energy increases at a faster rate than the average kinetic energy means that there is additional energy available for particle creation. In the present context this means that there is additional energy available that can be released in the form of seismic waves in excess to the energy which is converted into heat and displacement work.

In Section 1 we mentioned that aftershocks have the property that the ratio of the volume affected by an aftershock, V, to the earthquake volume of the main shock, $V^\star$, should have the same distribution as the ratio of the energy of an aftershock, ε to the energy of the main shock, $\varepsilon^\star$. Consequently, we may use the definition of the pressure P as $\partial S/\partial V = P/T$ to derive the mechanical equation of state

$$PV^{\star\,\tilde{\rho}}V^{1-\tilde{\rho}} = \left(\frac{\eta}{\eta^\star}\right)^{\tilde{\rho}}\mathrm{nT}. \tag{5.6}$$

This mechanical equation of state relates the positive definiteness of the isothermal compressibility, $\kappa = -(1/V)\,(\partial V/\partial P)_T = \{(1-\tilde{\rho})P\}^{-1}$ to the criterion of the concavity of the entropy, (5.1). At the extremes, $\tilde{\rho} = 0$ and $\tilde{\rho} = 1$, (5.6) would reduce to the mechanical equation of state of an ideal gas for the volume affected by an aftershock of energy ε and the earthquake volume of the main shock of energy $\varepsilon^\star$, respectively. That a power of the volume less than one appears in the mechanical equation of state may be indicative of a fractal structure, of either dimension $3(1-\tilde{\rho})$ or $3\tilde{\rho}$ embedded in an ambient dimension of 3, for the effective volume affected by an aftershock, or the main shock, respectively. Furthermore, (5.6) predicts that the two volumes behave in a complementary manner to one another.

Finally, when the temperature is eliminated between the thermal, (5.5), and mechanical, (5.6), equations of state there results the equation of state

$$P = \eta.$$

It states that the radiation pressure developed by the seismic waves of aftershocks is determined uniquely by the energy density, η. Likewise, the pressure of the main shock is $P^\star = \eta^\star$. This is analogous to the radiation pressure of electromagnetic radiation which is determined by the energy density in the black-body cavity.

Glossary

– *Extreme value distribution* – The limiting distribution for the largest or smallest element of a set of independent and identically distributed variables. If this distribution is of the exponential type then the distribution of largest value is the Gumbel distribution. If the distribution is an inverse power law then the distribution of the largest value is the Fréchet distribution, while if the distribution is a positive power law then the distribution of the smallest value is a Weibull distribution.

– *Hazard rate* – The hazard, or failure, rate is a nondecreasing convex function that is used in life testing; the conditional probability density of the lifetime of an item given that it has survived to a specific time. The hazard function of a Weibull distribution with exponent 2 is a straight line. The reciprocal of the hazard function of a Pareto distribution is also a straight line. The reciprocal hazard and mean excess functions are proportional to one another.

– *Mean excess function* – The mean excess function, or the mean residual life function, is the conditional expectation $X-x$ given that $X>x$. The mean excess function of an inverse power law, or Pareto, distribution is a straight line provided the exponent is greater than unity.

– *Order statistics* – Order statistics results when the random variables are arranged in order from the smallest to the largest, or *vice versa*. Notwithstanding the fact that the values so derived are no longer independent nor identically distributed, even though the original were, order statistics has some remarkably simple properties. In particular, the sampling distribution of the transformed order statistic is a beta distribution, and the joint distribution is a product of beta distributions, just as in the case of independent and identically distributed random variables.

– *Quantile function* – For strictly increasing and continuous distributions, the quantile function is the usual inverse of the distribution. The q-quantile x is a value along the measurement scale with the property that the fraction q of the distribution is left of x. The quantile function provides visual discrimination between distributions. Whereas both quantile functions of the Fréchet, $(-\ln(q))^{-1/\alpha}$, and Weibull, $-(-\ln(q))^{1/\alpha}$, distributions are increasing over the entire interval $q \in [0,1]$ for $\alpha > 0$, the Fréchet quantile function is convex while the Weibull quantile function has an inflexion point which separates a concave function for small values of q and a convex function for large values of q.

– *Return period* – If an event has probability p then, on the average, $1/p$ trials will have to be made in order that the event will happen once. The inverse of the probability is the return period, or the mean of the first exceedance time.

Acknowledgements

This work was supported, in part, by Programma Accordo MICA-ENEA, and MURST ex-40% and 60%.

REFERENCES

ABACUS CONCEPTS (1994): *Survival Tools for StatView* (Abacus Concepts).

ALDOUS, D. (1989): *Probability Approximations via the Poisson Clumping Heuristic* (Springer-Verlag), 44-45.

BÅTH, M. (1965): Lateral inhomogeneities of the upper mantle, *Tectonophysics*, **2**, 483-514.

BÅTH, M. and S.J. DUDA (1964): Earthquake volume, fault plane area, seismic energy, strain, deformation and related quantities, *Ann. Geofis.*, **17**, 353-368.

CRAMÉR, H. (1946): *Mathematical Methods of Statistics* (Princeton University Press), 370-371.

EPSTEIN, B. (1948): Statistical aspects of fracture problems., *J. Appl. Phys.*, **19**, 140-147.

EVISON, F. (1999): On the existence of earthquake precursors, *Ann. Geofis.*, **42** (5), 763-770.

DE FINETTI, B. (1970): *Theory of Probability* (Wiley), vol. 2, p. 102.

GUMBEL, E.J. (1958): *Extreme Statistics* (Columbia University Press), p. 151.

GUTENBERG, B. (1956): Earthquake magnitude, intensity, energy and acceleration, *Bull. Seismol. Soc. Am.*, **46**, 105-145.

HASTINGS, H.M. and G. SUGIHARA (1993): *Fractals: a User's Guide for the Natural Sciences* (Oxford University Press), 44-46.

KNOPOFF, L. and Y. KAGAN (1977): Analysis of the theory of extremes as applied to earthquake problems, *J. Geophys. Res.*, **82**, 5647-5657.

KURIMOTO, H. (1959): A statistical study of some aftershock problems, *Zisin*, **12**, 1-10.

LAVENDA, B.H. (1991): *Statistical Physics: a Probabilistic Approach* (Wiley-Interscience), p. 262.

LAVENDA, B.H. (1995): *Thermodynamics of Extremes* (Horwood), p. 45.

LAVENDA, B.H. and E. CIPOLLONE (2000): Extreme value statistics and thermodynamics of earthquakes: large earthquakes. *Ann. Geofis.*, **43** (3), 469-484.

LEADBETTER, R. and H. ROOTZEN (1988): Extremal theory for stochastic processes, *Ann. Prob.*, **16**, 431-478.

LOMNITZ, C. (1966): Magnitude stability in earthquake sequences, *Bull. Seismol. Soc. Am.*, **56**, 247-249.

LOMNITZ, C. (1994): *Fundamentals of Earthquake Prediction* (Wiley), p. 151.

LOMNITZ, C. and A. HAX (1966): Clustering of aftershock sequences, *The Earth Beneath the Continents*, Am. Geophys. Un., *Geophys. Monogr.* 10, 502-508.

LOMNITZ-ADLER, J. and C. LOMNITZ (1978): A new magnitude-frequency relation, *Tectonophysics*, **49**, 237-245.

OMORI, F. (1894): On the aftershocks of earthquakes, *J. Coll. Sci. Imp. Univ. Tokyo*, **7**, 111-200.

RANALLI, G. (1969): A statistical study of aftershock sequences, *Ann. Geofis.*, **22** (4), 359-397.

REISS, R.-D. and M. THOMAS (1997): *Statistical Analysis of Extreme Values* (Birkhäuser), p. 13.

RICHTER, C.F. (1958): *Elementary Seismology* (Freeman, San Francisco).

SOLOV'EV, S.L. and O.N. SOLOV'EVA (1969): Exponential increase of the total number of an earthquake's aftershocks and the decrease of their mean value with increasing depth, *Izv. Akad. Nauk. SSSR, Ser. Geofiz.*, **12**, 1695-1694, English translation 1053-1060.

STUART, A. and K. ORD (1994): *Kendall's Advanced Theory of Statistics*, vol. I, 6th edition (E. Arnold), p. 475.

UTSU, T. (1969-1970): Aftershocks and earthquake statistics (two papers), *J. Fac. Sci. Hokkaido U., Ser. VII*, **3**, 129-195 and 379-441.

UTSU, T., Y. OGATA and R. MATSU'URA (1995): The centenary of the Omori formula for the decay law of aftershock activity, *J. Phys. Earth*, **43**, 1-33.

VERE-JONES, D. (1966): A Markov model for aftershock occurrence, *Pure Appl. Geophys.*, **64**, 31-42.

VERE-JONES, D. (1969): A note on the statistical interpretation of Båth's law, *Bull. Seismol. Soc. Am.*, **59** (4), 1535-1541.

(received May 5, 2000;
accepted October 31, 2000)