ANALYSIS OF THE THEORY OF EXTREMES AS APPLIED TO EARTHQUAKE PROBLEMS

L. Knopoff and Y. Kagan

Institute of Geophysics and Planetary Physics,
University of California, Los Angeles, California 90024

Abstract. Procedures of the theory of extremes give unacceptably large probable errors in determinations of return times, b values, and maximum magnitudes of large- and intermediate-magnitude earthquakes using limited runs of data. On all accounts, methods which utilize all available data give superior estimates of the parameters of seismicity than do extreme value methods and provide for a procedure for handling inhomogeneous catalogs as well.

Introduction

The theory of extremes developed by Gumbel [1958] provides a convenient method to obtain estimates of the frequencies of occurrence of events on the extreme of a statistical distribution and to estimate recurrence times for these events, provided certain conditions involved in the theory are met. All too often the theory is applied without regard for the constraining conditions; in this paper we investigate the consequences of applying one of the constraints rigorously.


With regard to application of the theory of extremes to problems of earthquake occurrence and recurrence, Epstein and Lomnitz [1966] have shown that Gumbel’s statistical model I of the theory of extremes can be derived directly from the assumptions that earthquakes in a given catalog are produced by a simple Poisson process and obey the usual magnitude-frequency law. We note that Gumbel [1958] and Jenkinson [1955] have interchanged the order of their reference to models I, II, and III. Jenkinson has shown that the three asymptotic models are obtainable as a solution to a functional equation due to Fisher and Tippett [1928]. In Jenkinson’s solution the three models are represented as a function of the exponent k in a power law relationship.
magnitude formula demand that there be an upper bound to the magnitude of earthquakes [Newmark and Rosenblueth, 1971]. The demonstration is as follows: The usual cumulative frequency law is written as

\[ \log N = a - bM \]

and the energy-magnitude formula is

\[ \log E = a + bM \]

where \( N \) is the rate of occurrence of earthquakes with magnitudes greater than \( M \).

The rate of occurrence of earthquakes with energies between \( E \) and \( E + dE \) is

\[ dN = \text{const.} E^{-b/8 - 1} dE \]

and the total amount of energy released in earthquakes per unit time is

\[ E_{max} \frac{dE}{E_{min}} = \text{const.} E^{1 - b/8} \]

Typically, \( b \approx 1 \), and \( b/8 \approx 0.6 \). Thus \( E_{max} \) must be finite so that the mean rate of energy released is finite. (If \( b/8 \) were greater than 1, then \( E_{min} \) would have to be nonzero by the same reasoning).

While this demonstration shows that there must be a finite upper bound to the largest earthquake, if the usual magnitude-frequency law applies over the entire range of allowable magnitudes (Figure 1), it does not identify what this bound is. The fact that the largest earthquakes reported in the worldwide catalog have \( M = 8.9 \) or thereabouts [Duda, 1965] is in consequence of other, considerations, most likely of geometrical and rheological natures.

We can avoid the \( E_{max} \) catastrophe by requiring that the magnitude-frequency law have a different form. We might imagine \( b \) to be a function of \( M \) in such a way that \( b \) increases with increasing \( M \), so that the energy integral remains finite. In one version of this we might assume that the differential form of the usual magnitude-frequency law applies over the range \( M < M_{max} \); the cumulative magnitude frequency formula thus obtained is the lower curve of Figure 1.

The assumption that there is no lower bound to the detectability of earthquakes in a given area is unreasonable in view of the finiteness of seismological networks. Knopoff and Gardner [1969] have given an empirical method for determining the lower-magnitude threshold of reliable reporting by a network; for the southern California network it is \( M = 3 \) after 1952. For the U.S. Coast and Geodetic Survey catalog it is about \( M = 4.5 \) [Knopoff and Gardner, 1972]. If there is a lower observational magnitude threshold, there is a finite probability that in certain years the largest naturally occurring event will fall below the threshold magnitude, and hence the catalog will have no entry for those years. The distribution of extreme values for earthquakes in a given region cannot be constructed without certain arbitrary assumptions regarding the years without recorded data.

In the large number of papers cited above, these two constraints have been ignored, namely, that the ordinary magnitude-frequency relation cannot be extended to ever larger magnitudes indefinitely and that in some years no earthquake having a magnitude greater than the lower threshold will occur. The postulate made in these papers is that the assumption of Poissonian character and the form of the ordinary magnitude-frequency relation are sufficient to permit the reconstruction of the detailed magnitude-frequency relation and the estimate of the recurrence time and, in some cases, the size of the largest possible earthquakes in a given area from the observations of the largest annual events.

Synthesized Catalogs

To obtain an understanding of the reliability of extreme value analysis and especially to understand the importance of the maximum allowable earthquake we have synthesized hypothetical extreme value catalogs as follows: the usual cumulative magnitude-frequency law is

\[ \log_{10} N = -0.85(M - 4.0) + \log_{10} 20 = 4.70 - 0.85M = -0.85(M - 5.53), \quad M < \infty \]

This formula has a Poissonian rate of occurrence of earthquakes of magnitude 4 or greater of 20/yr. This rate, plus the value of \( b = 0.85 \), is not inappropriate for southern California [Richter, 1958]. The annual rate of occurrence of events with magnitudes greater than \( M \) is given in Table 1.

The differential version of the above formula

\[ (dN/N) \log_{10} e = 0.85 \frac{dM}{E_{max}} \]

is assumed to apply over the range of magnitude 4.0 \( \leq M \leq M_{max} \); the cumulative version of this relation is plotted in Figure 1. In
the numerical experiments we have used three
different values of the maximum magnitude:
M_{max} = 12.0, 8.0, 7.0. For each of these
cases, 15 realizations of a Poisson process
were made by a Monte Carlo procedure; each
realization corresponded to a sequence of 40
years' duration. Hence each of the 15
realizations had approximately 800 earthquakes
of magnitude 4.0 \leq M \leq M_{max}. Over a period
of 40 years of earthquakes and 15 realizations
per model we have generated 600 years of earth-
quakes numerically. The rate of recurrence of
earthquakes with M > M_{max} in a system un-
bounded at the upper energy limit, delineates
the components of the magnitude spectrum that
are lost by truncation at the upper limit.

From Table 1 we see that if we let M_{max} = 12,
for all practical purposes it is equivalent to
setting M_{max} = \infty. In a 600-year list for
M_{max} = 12, there is a probability of about \frac{1}{n}
that an earthquake with M > 9 will appear. In
the 600-year list for M_{max} = 9, there is a
probability of
0.984 \approx (0.992 - 0.491)/(1.000 - 0.491)
that an earthquake with 9 \geq M \geq 8 will appear.
The point of this discussion is that there is a
finite chance that a M > 9 event will appear
in a computer-generated list of earthquakes with
M_{max} = 12 if the list is of sufficient length.

We have synthesized a list of 40 years' duration in
order to simulate the catalog of southern California earthquake [Hileman et al., 1973], which starts essentially at the time of the
Long Beach earthquake of 1933. We indicate
below that scaling is possible in terms of the
ratio of return times to the number of
years of data so that the illustration present-
ed here can be extended to other spans of
catalogs.

Plotting Rules

For each of the 15 realizations for each of
the three values of M_{max}, the output was
processed in the usual way. The 40 extreme
annual earthquakes were rank-ordered. For the
purposes of display the magnitudes of the rank-
ordered annual earthquakes, for constant rank
order number, were averaged over all 15 reali-
izations; these averaged values are plotted on
probability paper as a function of the
number of years of data; in our case, n = 40.

Several versions of the plotting rule G(m)
have been proposed. Gumbel's [1958] plotting
rule is
\[ G(m) = \frac{m}{n + 1}, \]
where m is the rank order number and n is the
number of years of data; in our case, n = 40.
This plotting rule is favored by most
seismologists, perhaps because it has the
property that the return period of the
largest of the n annual samples is n years and
because a tradition descended from Gumbel
governs its use. The literal result for the case
M_{max} = 12 diverges significantly from the
theoretical curve for M_{max} = \infty, n = \infty (Figure 2).

<table>
<thead>
<tr>
<th>M</th>
<th>N, yr^-1</th>
<th>M^-1, yr</th>
<th>Probability of</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.0</td>
<td>3.2 x 10^-6</td>
<td>3.2 x 10^5</td>
<td>0.002</td>
</tr>
<tr>
<td>9.0</td>
<td>1.1 x 10^-3</td>
<td>889.</td>
<td>0.491</td>
</tr>
<tr>
<td>8.0</td>
<td>8.0 x 10^-3</td>
<td>126.</td>
<td>0.992</td>
</tr>
<tr>
<td>7.0</td>
<td>5.6 x 10^-2</td>
<td>17.7</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Thus there is significant bias in using plotting
rule (1).

The following heuristic argument can be
offered concerning the reason for the bias in
using plotting rule (1). Consider the refer-
ence earthquake whose magnitude is such that it
has a true recurrence rate of n years. If the
magnitude of the largest earthquake in a sample
of n annual earthquakes is less than that for
the reference earthquake, it is likely to be
relatively close to the reference magnitude;
whereas in this case the return time predicted
for the reference earthquake will be only
slightly larger than n years. However, if the
magnitude of the largest earthquake in a sample
of n annual events is greater than that for the
reference earthquake, it could be significantly
larger than the reference magnitude, a return
time of the reference earthquake thus being
predicted which is significantly less than n
years. In other words the distribution of the
largest of the n annual samples is skewed to
lower magnitudes than the reference value and
has a small tail extending to larger magnitudes.
The skew distribution results in a bias for
average return times for the largest of the
n annual samples which will predict lower
values than n years.

A confirmation of this point of view is
given by Gumbel [1958, pp. 113-116]. The return
period of the average largest value for an
exponential distribution, which is appropri-
ate in our case, is 1.78n, whereas the return
period of the most probable extreme value is
n. Thus in our case we should expect the re-
turn period for the average largest value of
our 40-year samples to be 1.78 x 40 = 71 years
and not 40 years as plotting rule (1) pre-
dicts. For the plotting rule
\[ G(m) = \frac{(m - A)/(n + 1 - 2A)}{n} \]
which has a common value for G(n + 1)/2 for
any value of A, we prefer to use a value of A
which is positive and nonzero in preference to
the value A = 0 used by Gumbel and his
successors.

It might be thought that the use of the
version of G(m) given by (1) might still be
acceptable on statistical grounds. Each of the
means of the 15 realizations has a certain
A third plotting rule is given by Gringorten [1963a] for the case
\[ A = 1 - e^{-\gamma} = 0.439 \]
where \( \gamma \), Euler's constant, is equal to 0.577...; i.e.,
\[ G(m) = \frac{m - 0.44}{n + 0.12} \quad (4) \]
This function minimizes the bias in the long-return period end of the distribution. For the case \( n = 40 \) we are not able to distinguish between the two plotting rules (3) and (4). In the results reported below, we have used the simple rule (3).

### Events Unbounded In Maximum Magnitude

We have remarked that the differences between the extreme values plotted in Figure 2 and the theoretical straight line are in general small except at the larger magnitudes. For smaller magnitudes the differences from the theoretical straight line are much less than the variations in the plots for the individual realizations. Not only do the individual realizations show a much greater scatter than is shown in Figure 2, but also it is often difficult to discriminate between a curve generated for \( M_{\text{max}} = 12 \) and one for \( M_{\text{max}} = 7 \).

If the assumption is made that \( M_{\text{max}} = \infty \), then according to Epstein and Lomnitz [1966] the parameters \( (a, b) \) of the frequency-magnitude formula can be recovered by linear regression. We have calculated the regression lines for each individual realization as well as for the mean curves of Figure 2 for each of the three postulates regarding \( M_{\text{max}} \). In the cases of \( M_{\text{max}} = 7.0 \) and 8.0, Figure 2 shows that the curves are not straight lines, but before we proceed to a more complex analysis, let us assume that we interpret these curves of Figure 1 as though \( M_{\text{max}} = \infty \). The detailed analysis for the case \( M_{\text{max}} = 12 \) is presented in Table 2; a summary of the regression analysis for the other two cases is shown in Table 3.

The parameter \( M_0 \) is the least squares estimate for \(-\log[-\log G(m)] = 0\), i.e., \( G(m) = e^{-t} \), corresponding to \( m \approx 15 \) for \( n = 40 \). The slope of the least squares line is \( (b \ln 10)^{-1} \). The values of \( b \) and \( M_0 \) are tabulated in Table 3 along with their uncertainties. The recurrence time for an earthquake with \( M > 8.5 \) is calculated from
\[ T_M > 8.5 = 10^{(8.5 - M_0) b} \text{ years} \]

For the purposes of comparison we have listed in the row 'regression of means' in Table 2 the result of regression analysis on the curve \( M_{\text{max}} = 12 \) shown in Figure 2, which is the average of the 15 rank-ordered magnitudes. It is seen that the error estimates of the values of \( (M_0, b) \) for the smoothed averaged curve are about what might be expected, namely, about a factor of \( (14)^{1/2} \), or 4 less than the error estimates for the individual realizations. However, we note that the scatter in the b values for the 15 individual realizations is much greater than the error estimate of
TABLE 2. 15 Monte Carlo Realizations of Earthquake Sequences with $M_{\text{max}} = 12$

<table>
<thead>
<tr>
<th>Realization number</th>
<th>$M_o$</th>
<th>$\sigma_M$</th>
<th>$b$</th>
<th>$\sigma_b$</th>
<th>$T_{M &gt; 8.5}$</th>
<th>$\sigma_T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.576</td>
<td>0.028</td>
<td>0.665</td>
<td>0.021</td>
<td>88</td>
<td>1.161</td>
</tr>
<tr>
<td>2</td>
<td>5.586</td>
<td>0.033</td>
<td>0.662</td>
<td>0.024</td>
<td>85</td>
<td>1.186</td>
</tr>
<tr>
<td>3</td>
<td>5.576</td>
<td>0.018</td>
<td>0.752</td>
<td>0.018</td>
<td>157</td>
<td>1.130</td>
</tr>
<tr>
<td>4</td>
<td>5.472</td>
<td>0.020</td>
<td>1.211</td>
<td>0.049</td>
<td>4648</td>
<td>1.416</td>
</tr>
<tr>
<td>5</td>
<td>5.312</td>
<td>0.017</td>
<td>0.932</td>
<td>0.025</td>
<td>937</td>
<td>1.209</td>
</tr>
<tr>
<td>6</td>
<td>5.467</td>
<td>0.014</td>
<td>1.186</td>
<td>0.034</td>
<td>3946</td>
<td>1.269</td>
</tr>
<tr>
<td>7</td>
<td>5.587</td>
<td>0.020</td>
<td>0.993</td>
<td>0.033</td>
<td>783</td>
<td>1.258</td>
</tr>
<tr>
<td>8</td>
<td>5.700</td>
<td>0.030</td>
<td>0.718</td>
<td>0.026</td>
<td>102</td>
<td>1.195</td>
</tr>
<tr>
<td>9</td>
<td>5.529</td>
<td>0.029</td>
<td>0.739</td>
<td>0.027</td>
<td>156</td>
<td>1.209</td>
</tr>
<tr>
<td>10</td>
<td>5.641</td>
<td>0.032</td>
<td>0.957</td>
<td>0.049</td>
<td>544</td>
<td>1.395</td>
</tr>
<tr>
<td>11</td>
<td>5.573</td>
<td>0.021</td>
<td>0.670</td>
<td>0.016</td>
<td>91</td>
<td>1.118</td>
</tr>
<tr>
<td>12</td>
<td>5.539</td>
<td>0.022</td>
<td>0.845</td>
<td>0.027</td>
<td>318</td>
<td>1.204</td>
</tr>
<tr>
<td>13</td>
<td>5.626</td>
<td>0.023</td>
<td>0.780</td>
<td>0.024</td>
<td>174</td>
<td>1.178</td>
</tr>
<tr>
<td>14</td>
<td>5.461</td>
<td>0.013</td>
<td>0.779</td>
<td>0.013</td>
<td>232</td>
<td>1.100</td>
</tr>
<tr>
<td>15 Regression of means</td>
<td>5.558</td>
<td>0.016</td>
<td>0.748</td>
<td>0.015</td>
<td>158</td>
<td>1.114</td>
</tr>
<tr>
<td>Statistics of realizations</td>
<td>5.547</td>
<td>0.003</td>
<td>0.813</td>
<td>0.003</td>
<td>250</td>
<td>1.022</td>
</tr>
<tr>
<td>Average of reciprocals of $b$</td>
<td>5.547</td>
<td>0.092</td>
<td>0.842</td>
<td>0.178</td>
<td>306</td>
<td>3.40</td>
</tr>
<tr>
<td>$M_{\text{max}} = \infty$</td>
<td>5.53</td>
<td>0.85</td>
<td></td>
<td></td>
<td></td>
<td>334</td>
</tr>
</tbody>
</table>

0.003 x (14)$^{1/2}$ to be expected from the smoothed curve of averages of Figure 2 or from the estimate of standard deviation for the individual realizations. An analysis of the scatter in the values of $(M_o, b, T)$ obtained from the 15 regression analyses gives the row labeled 'statistics of realizations.' In this case the scatter in $M_o$ is about 4 times larger than the errors for the individual realizations, and the scatter in $b$ is about 6 times larger. The values in parentheses are obtained by the averages of the reciprocals of $b$; this would seem to be a reasonable procedure, since $b$ is related to the reciprocal slope of the curve in Figure 2. However, a direct averaging of the $b$ values for each realization gives a closer approach to the $b$ value for the original model; the parameters for the original model, i.e., the expectation for $M_{\text{max}} = \infty$ with an infinite number of realizations, is given on the bottom line.

Since the return time is obtained by exponentiation, the mean values are described in terms of their geometric means rather than their arithmetic means, and the error estimates are indicated as multiplicative rather than additive factors. In the case of the estimate of the return time for an earthquake of magnitude greater than 8.5 we see that one realization gives a very optimistic estimate, i.e., to about 20% for one standard deviation. But the scatter in the return times among the various realizations is enormous, with a standard error of more than 3 times the estimate; we interpret the bounds for one standard deviation as $(3.40 \pm 3) \times 306$.
TABLE 3. Mean Regression for 15 Monte Carlo Realizations of Earthquake Sequences

<table>
<thead>
<tr>
<th>M_0</th>
<th>σ_M</th>
<th>b</th>
<th>σ_b</th>
<th>T_{M &gt; 8.5, yr}</th>
<th>Multiplicative Uncertainty ( σ_T )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( M_{\text{max}} = 8 )</td>
</tr>
<tr>
<td>Regression of Means</td>
<td>5.540</td>
<td>0.010</td>
<td>0.903</td>
<td>0.014</td>
<td>471</td>
</tr>
<tr>
<td>Statistics of Realizations</td>
<td>5.540</td>
<td>0.084</td>
<td>0.922</td>
<td>0.143</td>
<td>533</td>
</tr>
<tr>
<td>Average of Reciprocals of b</td>
<td></td>
<td></td>
<td>(0.903)</td>
<td>(0.133)</td>
<td>(470)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( M_{\text{max}} = 7 )</td>
</tr>
<tr>
<td>Regression of Means</td>
<td>5.513</td>
<td>0.018</td>
<td>1.080</td>
<td>0.035</td>
<td>1678</td>
</tr>
<tr>
<td>Statistics of Realizations</td>
<td>5.513</td>
<td>0.087</td>
<td>1.088</td>
<td>0.096</td>
<td>1777</td>
</tr>
<tr>
<td>Average of Reciprocals of b</td>
<td></td>
<td></td>
<td>(1.080)</td>
<td>(0.098)</td>
<td>(1682)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( M_{\text{max}} = \infty )</td>
</tr>
<tr>
<td>Theoretical, ( n = \infty )</td>
<td>5.53</td>
<td>0</td>
<td>0.850</td>
<td>0</td>
<td>334</td>
</tr>
</tbody>
</table>

years. This indicates that the estimate of return times for larger earthquakes, based on one realization, is worthless; of course in nature we only have one realization. We note that a similar calculation using the Gumbel rule (1) for \( G(m) \) would have given a significant underestimate of \( b \), namely, 0.774 ± 0.159, and an underestimate of the return time \( T = 199 \) years with ±1 standard deviation times at 62 and 644 years, i.e., a factor of 3.24. These results confirm the better estimate to be obtained by using the definition \( (m - \bar{x})/\bar{n} \) for probability.

By any measure of analysis the fluctuations in the return times and the \( b \) values are enormous; this means that the derivation of \( b \) or return times from one realization, the natural one, is subject to very great uncertainties, and such results can only be considered as very unreliable.

It is clear that the large uncertainty in estimating the return time of an earthquake of magnitude 8.5 is due to the enormous extrapolation from 40 years of data to predict an event with a return time of 334 years. We expect that the uncertainties in the estimates of the return times of earthquakes with lower magnitudes will be smaller. Inspection of Table 4 shows that the estimates of return times have large fractional uncertainties even for relatively low magnitudes. For example, the return time for an earthquake of magnitude 7.15 can be estimated only to within a factor of 2 at the 10 level, or a factor of 4 at the 20 level; the return time of such an earthquake is 23.8 years, a time interval rather smaller than the 40 years of data collection. Thus the results of regression analysis on a set of real, naturally derived data lead to estimates of return times that are subject to large probable errors down to relatively low magnitudes; of course the probable errors diminish with increasing length of unbroken historical record. The last column of Table 4 provides a measure of the size of an earthquake: the appropriate scaling factor is in terms of the ratio of the return time to the duration of the earthquake sequence.

As we have noted, we have followed the suggestion of Epstein and Lomnitz [1966] that least squares regression analysis be applied. The use of least squares method for this problem is most questionable. Least squares can be applied if errors in the original data are independent and normally distributed. But the points in the extreme values plot are not independent, since the curves are monotonically increasing: each subsequent value is greater than or equal to the preceding one. This fact explains why the estimates of errors, obtained from each individual realization of the process (Table 2) are about 4-6 times less than the real scatter of the estimates of parameters. Thus the standard errors obtained in the least squares fit of the extreme values plot may be misleading. In addition, the hypothesis of normality cannot be satisfied at all, since the lower part of the distribution for each point is cut off at the value of the preceding point. Better estimates of the parameters are provided
by the maximum likelihood method, but its application to the extreme values problem is very complicated [Gumbel, 1958, p. 231].

Analysis for Maximum Magnitudes

The curvature in the displays of Figure 2 for $M_{\text{max}} = 8.0$ and $7.0$ suggests that one might try Gumbel's model III to see if the magnitude-frequency parameter as well as the maximum size of the earthquake can be estimated. Makjanic [1972] and Yegulalp and Kuo [1974] propose that earthquake statistics indicate a maximum size; this proposal was made on the basis of curvature of the graph of annual extreme values and not on any theoretical grounds. The model is

$$M - M_o = a \ln 1 - \exp \left( -k (\ln G(m)) \right)$$

When $k = 0$, $a$ finite, we have the degenerate case that we have considered up to this point. If $k > 0$, then there will be a maximum possible earthquake; the value of $b$ is $(4k \ln 10)^{-1}$. We have no compelling physical model that indicates that the extreme value statistic for the truncated magnitude-frequency relation is Gumbel's model III for $n$ finite. Nevertheless, we can try to fit a curve of this type to the realizations to estimate the three parameters ($b$, $M_{\text{max}}$, $M_o$) by least squares. The least squares procedure for Gumbel's model III is given by Jenkinson [1955] Since most of the entries are in the lower magnitude part of the range, and hence in the linear part of the curve, the accuracy with which the quantity $M_{\text{max}}$ might be obtained will be poor.

The results of least squares regression analysis applied to each of the fifteen realizations for the case $M_{\text{max}} = 7.0$ are tabulated in Table 5. (As in the case of model I, the values of $M_o$, and $(b \ln 10)^{-1}$ are the intercept and slope on the axis $G(m) = 1/e$.) It can be seen that the values of $M_o$ and $b$ are in general close to the correct values; however, the prediction of the maximum magnitude has proved to be highly unstable, estimates ranging from $M_{\text{max}} = 7.1$ to $9.0$ with a large uncertainty about the mean of $M_{\text{max}} = 8.2$ (the correct answer is, of course, $7.0$). We have repeated the calculation with two additional sets of 15 realizations, and estimates of $M_{\text{max}}$ change from 7.0 in both cases to 10.6 and 13.7. A three-parameter regression analysis of the curve of Figure 2, which is the mean of the magnitudes at each rank-ordered probability value, gives comparable results. The huge scatter in estimates of $M_{\text{max}}$ indicates that 40 years of data are insufficient to determine this parameter.

With this background we may expect even greater instability in the three-parameter analysis of the case $M_{\text{max}} = 8.0$; with 40 years of data in each run the occurrence of an earthquake with magnitude close to 8.0 is very rare. But such large earthquakes are needed to define the bending of the magnitude-probability curves. Our expectations are wholly satisfied. In fact, for three of the 15 realizations, between two and four of the curves have a reversed curvature and hence resemble distributions of Gumbel type II more closely than they do type III. This implies a lower rather than an upper magnitude cutoff, and hence no $M_{\text{max}}$ can be estimated from these realizations. The regression analysis of the curve of Figure 2 gives $M_{\text{max}} = 10.7$.

Maximum Likelihood Estimates

The results that we have obtained indicate that the application of the theory of extremes leads to extraordinary instability in prediction of return times for large- and intermediate-magnitude earthquakes and in estimation of maximum magnitudes and even $b$ values. From the examples given, the difficulty can be directly attributed to insufficient data: in the attempt to predict the return of a $M = 8.5$ earthquake
the problem is to determine the theoretical return time of 334 years with the use of only 40 years of data. For the southern California magnitude-frequency formula an earthquake with a magnitude 7.4 has a return time of about 40 years. Thus the data from 40 years of accumulated annual earthquakes can be used to predict return times for earthquakes with $M > 7.4$, but only with extreme caution; the threshold is more appropriately considerably less than this.

It is hardly surprising that large uncertainties arise in extreme value estimates of the three parameters of $b$ value, return time, and maximum magnitude in comparison with such estimates obtained from the full catalog; the full catalog contains more information than its abstract of annual earthquakes. Thus we may assume that the extreme value calculation yields no information additional to that which can be obtained from analysis of the full catalog if the latter is available.

Is there any alternative statistical model which in fact is better than extreme value methods for determining the parameter of seismicity and yet gives some consolation to those who do not have access to a full catalog of seismicity? We propose that a full use of all the available data, rather than of the extreme value extract, still provides better estimates of the seismicity of a region, through the use of maximum likelihood methods. The application of the maximum likelihood method to the determination of the parameters in the magnitude-frequency relation has already been proposed by Aki [1965] and Kantorovich et al. [1970] and is a relatively simple procedure. The use of all the available data for the computation of the magnitude-frequency relation gives estimates with much lower scatter in comparison with estimates by extreme value methods, since the number of data must perform be greater in the maximum likelihood application. For example, the estimate of $b$ in Table 2 has a coefficient

<table>
<thead>
<tr>
<th>Realization number</th>
<th>$M_{\text{max}}$</th>
<th>$b$</th>
<th>$M_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.0</td>
<td>0.928</td>
<td>5.49</td>
</tr>
<tr>
<td>2</td>
<td>8.5</td>
<td>0.917</td>
<td>5.50</td>
</tr>
<tr>
<td>3</td>
<td>8.7</td>
<td>0.840</td>
<td>5.53</td>
</tr>
<tr>
<td>4</td>
<td>7.1</td>
<td>0.850</td>
<td>5.49</td>
</tr>
<tr>
<td>5</td>
<td>8.8</td>
<td>1.057</td>
<td>5.34</td>
</tr>
<tr>
<td>6</td>
<td>8.3</td>
<td>1.025</td>
<td>5.49</td>
</tr>
<tr>
<td>7</td>
<td>8.8</td>
<td>0.871</td>
<td>5.56</td>
</tr>
<tr>
<td>8</td>
<td>7.0</td>
<td>0.944</td>
<td>5.73</td>
</tr>
<tr>
<td>9</td>
<td>8.3</td>
<td>0.895</td>
<td>5.55</td>
</tr>
<tr>
<td>10</td>
<td>7.7</td>
<td>0.766</td>
<td>5.68</td>
</tr>
<tr>
<td>11</td>
<td>7.7</td>
<td>0.856</td>
<td>5.61</td>
</tr>
<tr>
<td>12</td>
<td>7.8</td>
<td>0.889</td>
<td>5.57</td>
</tr>
<tr>
<td>13</td>
<td>8.2</td>
<td>0.703</td>
<td>5.62</td>
</tr>
<tr>
<td>14</td>
<td>8.6</td>
<td>0.919</td>
<td>5.46</td>
</tr>
<tr>
<td>15</td>
<td>9.0</td>
<td>1.009</td>
<td>5.55</td>
</tr>
<tr>
<td>Regression of Means</td>
<td>7.9</td>
<td>0.891</td>
<td>5.55</td>
</tr>
</tbody>
</table>

Statistics of Realizations

Table 2

Theoretical

$7.0 \pm 0.85$ $5.53$
we may write a formula for the logarithm of
\[ m = (M - 3.9)/AM, \]
where \( AM \) is the magnitude maximum likelihood analysis. It is convenient
here to use the integer magnitude 
\[ m = (M - 3.9)/AM, \]
whence the maximum likelihood histogram (which is taken to be \( 0.1 \) in this case). Then using the Poissonian property of the model, we may write a formula for the logarithm of the likelihood function as [cf. Kantorovich, et al., 1971; Kagan and Knopoff, 1976]
\[
\ln L = \frac{1}{\gamma} \ln (\frac{M^r - M_{co}}{\gamma}) - \gamma - m_{max} + \sum_{m=1}^{m_{max}} \ln n_m \left( \gamma \cdot \frac{m_{co} - m}{\gamma} \right)
\]
In this formula, \( n_m \) is the number of earthquakes in the \( i \)th of the four parts of the catalog with integer magnitude \( m \), \( M_{co} \) is the lower cutoff magnitude in the \( i \)th part of the catalog, and \( m_{max} \) is a reference magnitude chosen to ensure minimum correlation between the estimates of \( \alpha \) and \( \gamma \) and is taken to be 5.85 in our case. The value of \( \alpha \) is the number of earthquakes in the interval \( M \), about \( m \) for the 40-year span of the catalog, i.e., within \( 5.9 > M \geq 5.8 \). The parameters \( (\alpha, \gamma) \)
correspond to \( (a, b) \) in the magnitude-frequency relation. In fact, \( \gamma = 10^{b/10} \).

A plot of \( \ln L \) as a function of \( \alpha \) and \( \gamma \) gives the maximum likelihood estimates of these parameters as well as the errors in these estimates.

The results of the maximum likelihood analysis of the truncated catalog are shown in Table 7. Although we are only using a little more than 30 earthquakes per subcatalog, the accuracy of the estimates of the \( a, b \) values and the return times is about the same as that of the extreme value analysis of the full catalog with about 40 annual entries; we recall that the extreme value analysis of the truncated catalog would have used only about 10 annual entries, and hence the errors in the estimates of \( b \) values and return times would have been much larger than the results of Table 2.

In the present case the estimates of the errors obtained from one realization correspond to the actual uncertainty. The return time is
\[
T_M = \frac{b^{0.85} \ln 10}{\alpha}, 10^{b(8.5 - 5.85)}
\]
where \( T \) is the time span of the catalog (40 years in this case). As we have noted, if we were to apply extreme value methods to these truncated data sets, we would either have to face the difficulty of coping with the missing data and the variation in the cutoff magnitudes for most of the early years or have to restrict our analysis to only the last 10 consecutive years of each example, with corresponding significant increase in uncertainties over those reported in Table 2. We may conclude that maximum likelihood methods yield estimates of the parameters of seismicity superior to those determined by extreme value methods, whether the catalogs are homogeneous or inhomogeneous.

Conclusions

1. Extreme value procedures using limited runs of data are unsatisfactory for the determination of \( b \) values, maximum magnitudes, and return times. Extreme value methods are unreliable for the estimation of return times greater than about one-half the span of the catalog. Most of the papers cited in the intro-

<table>
<thead>
<tr>
<th>Years</th>
<th>Cutoff Magnitude</th>
<th>Expected Number of Shocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-9.99</td>
<td>8.0</td>
<td>0.080</td>
</tr>
<tr>
<td>10.0-29.99</td>
<td>7.0</td>
<td>0.56</td>
</tr>
<tr>
<td>20.0-29.99</td>
<td>6.0</td>
<td>4.0</td>
</tr>
<tr>
<td>30.0-39.99</td>
<td>5.0</td>
<td>28.2</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>32.8</td>
</tr>
</tbody>
</table>

TABLE 6. Expected Numbers of Earthquakes

In 40-year catalog
### TABLE 7. Maximum Likelihood Estimates of Parameters

<table>
<thead>
<tr>
<th>Realization number</th>
<th>Number of Shocks in Subcatalogs</th>
<th>$5.9 &gt; M &gt; 5.8$</th>
<th>$\sigma_a$</th>
<th>$b$</th>
<th>$\sigma_b$</th>
<th>$T_M &gt; 8.5$</th>
<th>Multiplicative Uncertainty $\sigma_T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>35</td>
<td>4.46</td>
<td>0.62</td>
<td>0.821</td>
<td>0.106</td>
<td>254</td>
<td>2.20</td>
</tr>
<tr>
<td>2</td>
<td>37</td>
<td>4.87</td>
<td>0.75</td>
<td>0.737</td>
<td>0.088</td>
<td>125</td>
<td>1.96</td>
</tr>
<tr>
<td>3</td>
<td>29</td>
<td>3.68</td>
<td>0.70</td>
<td>0.906</td>
<td>0.126</td>
<td>570</td>
<td>2.53</td>
</tr>
<tr>
<td>4</td>
<td>36</td>
<td>4.12</td>
<td>0.75</td>
<td>1.055</td>
<td>0.137</td>
<td>1473</td>
<td>2.67</td>
</tr>
<tr>
<td>5</td>
<td>28</td>
<td>2.76</td>
<td>0.71</td>
<td>1.179</td>
<td>0.216</td>
<td>5238</td>
<td>4.59</td>
</tr>
<tr>
<td>6</td>
<td>34</td>
<td>3.92</td>
<td>0.76</td>
<td>1.011</td>
<td>0.188</td>
<td>1134</td>
<td>3.85</td>
</tr>
<tr>
<td>7</td>
<td>33</td>
<td>4.19</td>
<td>0.74</td>
<td>0.917</td>
<td>0.121</td>
<td>542</td>
<td>2.43</td>
</tr>
<tr>
<td>8</td>
<td>35</td>
<td>4.46</td>
<td>0.71</td>
<td>0.799</td>
<td>0.156</td>
<td>216</td>
<td>3.18</td>
</tr>
<tr>
<td>9</td>
<td>40</td>
<td>5.08</td>
<td>0.71</td>
<td>0.785</td>
<td>0.091</td>
<td>171</td>
<td>1.98</td>
</tr>
<tr>
<td>10</td>
<td>40</td>
<td>5.08</td>
<td>0.79</td>
<td>0.835</td>
<td>0.100</td>
<td>247</td>
<td>2.11</td>
</tr>
<tr>
<td>11</td>
<td>25</td>
<td>3.28</td>
<td>0.64</td>
<td>0.708</td>
<td>0.105</td>
<td>149</td>
<td>2.25</td>
</tr>
<tr>
<td>12</td>
<td>36</td>
<td>4.60</td>
<td>0.75</td>
<td>0.867</td>
<td>0.108</td>
<td>344</td>
<td>2.23</td>
</tr>
<tr>
<td>13</td>
<td>26</td>
<td>3.40</td>
<td>0.63</td>
<td>0.748</td>
<td>0.107</td>
<td>194</td>
<td>2.26</td>
</tr>
<tr>
<td>14</td>
<td>30</td>
<td>2.97</td>
<td>0.69</td>
<td>1.119</td>
<td>0.164</td>
<td>3204</td>
<td>3.22</td>
</tr>
<tr>
<td>15</td>
<td>32</td>
<td>4.24</td>
<td>0.71</td>
<td>0.726</td>
<td>0.093</td>
<td>132</td>
<td>2.05</td>
</tr>
</tbody>
</table>

Chaining of 15 realizations into one 600-year sample

<table>
<thead>
<tr>
<th>Statistics of realizations</th>
<th>Theoretical</th>
</tr>
</thead>
<tbody>
<tr>
<td>$33.1 \pm 4.7$</td>
<td>4.07</td>
</tr>
<tr>
<td>$0.73 \pm 0.149$</td>
<td>0.850</td>
</tr>
</tbody>
</table>

duction give results which very likely have unacceptably large uncertainties associated with them. Calculations of seismic risk do not provide estimates of the uncertainties should be viewed with skepticism.

2. Maximum likelihood methods are significantly superior to extreme value methods for the determination of the parameters of seismicity such as $b$ values and return times for homogeneous catalogs and provide a suitable method for deriving these parameters for inhomogeneous catalogs.

3. Monte Carlo procedures are useful methods for testing statistical procedures and are much to be preferred to those methods which test the internal consistency of a single sample.

4. Gumbel's plotting rule is significantly biased for large-magnitude events.

Acknowledgements. We are most grateful to David M. Perkins for his very constructive comments and suggestions, especially with regard to the questions of plotting rules. John K. Gardner gave important programming assistance to parts of this work. This research was supported by grant ENV 76-01706 of the National Science Foundation. Publication 1635 of the

Jenkinson, A. F., The frequency distribution of

Hald, A., Statistical Theory with Engineering

Hileman, J. A., C. R. Allen, and J. M. Nordquist,


Gringorten, I. I., Envelopes for ordered

Gringorten, I. I., A plotting rule for extreme


Epstein, B., and C. Lomnitz, A model for the

Duda, S. J., Secular seismic energy release in

Dick, I. D., Extreme value theory and earth-

Cox, D. R., and P. A. W. Lewis, The Statisti-

Court, A., Some new statistical techniques in

Aki, A., Maximum likelihood estimate of b in

Chen, P.-S., and P.-H. Lin, An application of

References

Aki, A., Maximum likelihood estimate of b in

Barlow, J. R., and F. J. Proschan, Mathematical

Borgan, O., A note on Frechet processes and

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhattacharya, R. N., A note on Frechet processes

Bhatcha