WHAT IS THE CHANCE OF AN EARTHQUAKE?

D. A. FREEDMAN AND P. B. STARK
Department of Statistics
University of California
Berkeley, CA 94720-3860

Technical Report 611.
September 2001; Revised January 2003.

1. Introduction

What is the chance that an earthquake of magnitude 6.7 or greater will occur before the year 2030 in the San Francisco Bay Area? The U.S. Geological Survey estimated the chance to be 0.7 ± 0.1 (*USGS*, 1999). In this paper, we try to interpret such probabilities.

Making sense of earthquake forecasts is surprisingly difficult. In part, this is because the forecasts are based on a complicated mixture of geological maps, rules of thumb, expert opinion, physical models, stochastic models, numerical simulations, as well as geodetic, seismic, and paleoseismic data. Even the concept of probability is hard to define in this context. We examine the problems in applying standard definitions of probability to earthquakes, taking the USGS forecast—the product of a particularly careful and ambitious study—as our lead example. The issues are general, and concern the interpretation more than the numerical values. Despite the work involved in the USGS forecast, their probability estimate is shaky, as is the uncertainty estimate.

2. Interpreting probability

Probability has two aspects. There is a formal mathematical theory, axiomatized by *Kolmogorov* (1956). And there is an informal theory that connects the mathematics to the world, i.e., defines what 'probability' means when applied to real events. It helps to start by thinking about simple cases. For example, consider tossing a coin. What does it mean to say that the chance of heads is 1/2? In this section, we sketch some of the interpretations—symmetry, relative frequency, and strength of belief.[1] We examine whether the interpretation of weather forecasts can be adapted for earthquakes. Finally, we present Kolmogorov's axioms and discuss a model-based interpretation of probability, which seems the most promising.

2.1. SYMMETRY AND EQUALLY LIKELY OUTCOMES

Perhaps the earliest interpretation of probability is in terms of 'equally likely outcomes', an approach that comes from the study of gambling. If the n possible outcomes of a chance experiment are judged equally likely—for instance, on the basis of symmetry—each must have probability $1/n$. For example, if a coin is tossed, $n = 2$; the chance of heads is $1/2$, as is the chance of tails. Similarly, when a fair die is thrown, the six possible outcomes are equally likely. However, if the die is loaded, this argument does not apply. There are also more subtle difficulties. For example, if two dice are thrown, the total number of spots can be anything from 2 through 12—but these eleven outcomes are far from equally likely. In earthquake forecasting, there is no obvious symmetry to exploit. We therefore need a different theory of probability to make sense of earthquake forecasts.

2.2. THE FREQUENTIST APPROACH

The probability of an event is often defined as the limit of the relative frequency with which the event occurs, in repeated trials under the same conditions. According to frequentists, if we toss a coin repeatedly under the same conditions,[2] the fraction of tosses that result in heads will converge to $1/2$: that is why the chance of heads is $1/2$. The frequentist approach is inadequate for interpreting earthquake forecasts. Indeed, to interpret the USGS forecast for the Bay Area using the frequency theory, we would need to imagine repeating the years 2000–2030 over and over again—a tall order, even for the most gifted imagination.

2.3. THE BAYESIAN APPROACH

According to Bayesians, probability means degree of belief. This is measured on a scale running from 0 to 1. An impossible event has probability 0; the probability of an event that is sure to happen equals 1. Different observers need not have the same beliefs, and differences among observers do not imply that anyone is wrong.

The Bayesian approach, despite its virtues, changes the topic. For Bayesians, probability is a summary of an opinion, not something inherent in the system being studied.[3] If the USGS says 'there is chance 0.7 of at least one earthquake with magnitude 6.7 or greater in the Bay Area between 2000 and 2030', the USGS is merely reporting its corporate state of mind, and may not be saying anything about tectonics and seismicity. More generally, it is not clear why one observer should care about the opinion of another. The Bayesian approach therefore seems to be inadequate for interpreting earthquake forecasts. For a more general discussion of the Bayesian and frequentist approaches, see *Freedman* (1995).

2.4. THE PRINCIPLE OF INSUFFICIENT REASON

Bayesians—and frequentists who should know better—often make probability assignments using Laplace's principle of insufficient reason (*Hartigan*, 1983, p. 2): if there is no reason to believe that outcomes are not equally likely, take them to be equally likely. However, not believed to be unequal is one thing; known to be equal is another. Moreover, all outcomes cannot be equally likely, so Laplace's prescription is ambiguous.

An example from thermodynamics illustrates the problem (*Feller*, 1968; *Reif*, 1965). Consider a gas that consists of n particles, each of which can be in any of r quantum states.[4] The state of the gas is defined by a 'state vector'. We describe three conventional models for such a gas, which differ only in the way the state vector is defined. Each model takes all possible values of the state vector—as defined in that model—to be equally likely.

1. Maxwell-Boltzman. The state vector specifies the quantum state of each particle; there are

$$r^n$$

possible values of the state vector.

2. Bose-Einstein. The state vector specifies the number of particles in each quantum state. There are

$$\binom{n+r-1}{n}$$

possible values of the state vector.[5]

3. Fermi-Dirac. As with Bose-Einstein statistics, the state vector specifies the number of particles in each quantum state, but no two particles can be in the same state. There are

$$\binom{r}{n}$$

possible values of the state vector.[6]

Maxwell-Boltzman statistics are widely applicable in probability theory,[7] but describe no known gas. Bose-Einstein statistics describe the thermodynamic behavior of bosons—particles whose spin angular momentum is an integer multiple of $\hbar$, Planck's constant h divided by 2π. Photons and He^4 atoms are bosons. Fermi-Dirac statistics describe the behavior of fermions, particles whose spin angular momentum is a half-integer multiple of $\hbar$. Electrons and He^3 atoms are fermions.[8]

Bose-Einstein condensates—very low temperature gases in which all the atoms are in the same quantum state—were first observed experimentally by *Anderson et al.* (1995). Such condensates occur for bosons, not fermions—compelling evidence for the difference in thermodynamic statistics. The principle of insufficient reason is not a sufficient basis for physics: it does not tell us when to use one model rather than another. Generally, the outcomes of an experiment can be defined in quite different ways, and it will seldom be clear a priori which set of outcomes—if any—obeys Laplace's dictum of equal likelihood.

2.5. EARTHQUAKE FORECASTS AND WEATHER FORECASTS

Earthquake forecasts look similar in many ways to weather forecasts, so we might look to meteorology for guidance. How do meteorologists interpret statements like 'the chance of rain tomorrow is 0.7'? The standard interpretation applies frequentist ideas to forecasts. In this view, the chance of rain tomorrow is 0.7 means that 70% of such forecasts are followed by rain the next day.

Whatever the merits of this view, meteorology differs from earthquake prediction in a critical respect. Large regional earthquakes are rare; they have recurrence times on the order of hundreds of years.[9] Weather forecasters have a much shorter time horizon. Therefore, weather prediction does not seem like a good analogue for earthquake prediction.

2.6. MATHEMATICAL PROBABILITY: KOLMOGOROV'S AXIOMS

For most statisticians, Kolmogorov's axioms are the basis for probability theory—no matter how the probabilities are to be interpreted. Let Σ be a σ-algebra[10] of subsets of a set S. Let P be a real-valued function on Σ. Then P is a probability if it satisfies the following axioms:

- $P(A) \geq 0$ for every $A \in \Sigma$;
- $P(S) = 1$;
- if $A_j \in \Sigma$ for $j = 1, 2, \ldots$, and $A_j \cap A_k = \emptyset$ whenever $j \neq k$, then

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P(A_j). \tag{1}$$

The first axiom says that probability is nonnegative. The second defines the scale: probability 1 means certainty. The third says that if A_1, A_2, $\ldots$

are pairwise disjoint, the probability that at least one A_j occurs is the sum of their probabilities.

2.7. PROBABILITY MODELS

Another interpretation of probability seems more useful for making sense of earthquake predictions: probability is just a property of a mathematical model intended to describe some features of the natural world. For the model to be useful, it must be shown to be in good correspondence with the system it describes. That is where the science comes in.

Here is a description of coin-tossing that illustrates the model-based approach. A coin will be tossed n times. There are 2^n possible sequences of heads and tails. In the mathematical model, those sequences are taken to be equally likely: each has probability $1/2^n$, corresponding to probability $1/2$ of heads on each toss and independence among the tosses.

This model has observational consequences that can be used to test its validity. For example, the probability distribution of the total number X of heads in n tosses is binomial:

$$P(X = k) = \binom{n}{k} \frac{1}{2^n}.$$

If the model is correct, when n is at all large we should see around $n/2$ heads, with an error on the order of $\sqrt{n}$. Similarly, the model gives probability distributions for the number of runs, their lengths, and so forth, which can be checked against data. The predictions hold quite well for real coins when the number of tosses is a few thousand. For hundreds of thousands of tosses, real coins do not behave exactly as the model predicts: differences are statistically significant.

This interpretation—that probability is a property of a mathematical model and has meaning for the world only by analogy—seems the most appropriate for earthquake prediction. To apply the interpretation, one posits a stochastic model for earthquakes in a given region, and interprets a number calculated from the model to be the probability of an earthquake in some time interval. The problem in earthquake forecasts is that the models—unlike the models for coin-tossing—have not been tested against relevant data. Indeed, the models cannot be tested on a human time scale, so there is little reason to believe the probability estimates. As we shall see in the next section, although some parts of the earthquake models are constrained by the laws of physics, many steps involve extrapolating rules of thumb far beyond the data they summarize; other steps rely on expert judgment separate from any data; still other steps rely on ad hoc decisions made as much for convenience as for scientific relevance.

3. The USGS Earthquake Forecast

We turn to the USGS forecast for the San Francisco Bay Area (*USGS*, 1999). The forecast was constructed in two stages. The first stage built a collection of 2,000 models for linked fault segments, consistent with regional tectonic slip constraints, in order to estimate seismicity rates. The models were drawn by Monte Carlo from a probability distribution defined using data and expert opinion.[11] We had trouble understanding the details, but believe that the models differed in the geometry and dimensions of fault segments, the fraction of slip released aseismically on each fault segment, the relative frequencies with which different combinations of fault segments rupture together, the relationship between fault area and earthquake size, and so forth.

Each model generated by the Monte Carlo was used to predict the regional rate of tectonic deformation; if the predicted deformation was not close enough to the measured rate of deformation, the model was discarded.[12] This was repeated until 2,000 models met the constraints. That set of models was used to estimate the long-term recurrence rate of earthquakes of different sizes, and to estimate the uncertainties of those rate estimates, for use in the second stage.

The second stage of the procedure created three generic stochastic models for fault segment ruptures, estimating parameters in those models from the long-term recurrence rates developed in the first stage. The stochastic models were then used to estimate the probability that there will be at least one magnitude 6.7 or greater earthquake by 2030.

We shall try to enumerate the major steps in the first stage—the construction of the 2,000 models—to indicate the complexity.

1. Determine regional constraints on aggregate fault motions from geodetic measurements.
2. Map faults and fault segments; identify fault segments with slip rates of at least 1 mm/y. Estimate the slip on each fault segment principally from paleoseismic data, occasionally augmented by geodetic and other data. Determine (by expert opinion) for each segment a 'slip factor', the extent to which long-term slip on the segment is accommodated aseismically. Represent uncertainty in fault segment lengths, widths, and slip factors as independent Gaussian random variables with mean 0.[13] Draw a set of fault segment dimensions and slip factors at random from that probability distribution.
3. Identify (by expert opinion) ways in which segments of each fault can rupture separately and together.[14] Each such combination of segments is a 'seismic source'.

4. Determine (by expert opinion) the extent to which long-term fault slip is accommodated by rupture of each combination of segments for each fault.
5. Choose at random (with probabilities of 0.2, 0.2, and 0.6 respectively) one of three generic relationships between fault area and moment release to characterize magnitudes of events that each combination of fault segments supports. Represent the uncertainty in the generic relationship as Gaussian with zero mean and standard deviation 0.12, independent of fault area.[15]
6. Using the chosen relationship and the assumed probability distribution for its parameters, determine a mean event magnitude for each seismic source by Monte Carlo simulation.
7. Combine seismic sources along each fault 'in such a way as to honor their relative likelihood as specified by the expert groups' (*USGS*, 1999, p. 10); adjust the relative frequencies of events on each source so that every fault segment matches its geologic slip rate—as estimated previously from paleoseismic and geodetic data. Discard the combination of sources if it violates a regional slip constraint.
8. Repeat the previous steps until 2,000 regional models meet the slip constraint. Treat the 2,000 models as equally likely for the purpose of estimating magnitudes, rates, and uncertainties.
9. Steps 1-8 model events on seven identified fault systems, but there are background events not associated with those faults. Estimate the background rate of seismicity as follows. Use an (unspecified) Bayesian procedure to categorize historical events from three catalogs either as associated or not associated with the seven fault systems. Fit a generic Gutenberg-Richter magnitude-frequency relation $N(M) = 10^{a-bM}$ to the events deemed not to be associated with the seven fault systems. Model this background seismicity as a marked Poisson process. Extrapolate the Poisson model to $M \geq 6.7$, which gives a probability of 0.09 of at least one event.[16]

This first stage in the USGS procedure generates 2,000 models and estimates long-term seismicity rates as a function of magnitude for each seismic source. We now describe the second stage—the earthquake forecast itself. Our description is sketchy because we had trouble understanding the details from the USGS report. The second stage fits three types of stochastic models for earthquake recurrence—Poisson, Brownian passage time (*Ellsworth et al.*, 1998), and 'time-predictable'—to the long-term seismicity rates estimated in the first stage.[17] Ultimately, those stochastic models are combined to estimate the probability of a large earthquake.

The Poisson and Brownian passage time models were used to estimate the probability that an earthquake will rupture each fault segment. Some parameters of the Brownian passage time model were fitted to the data, and some were set more arbitrarily; for example, aperiodicity (standard deviation of recurrence time, divided by expected recurrence time) was set to three different values, 0.3, 0.5, and 0.7. The Poisson model does not require an estimate of the date of last rupture of each segment, but the Brownian passage time model does; those dates were estimated from the historical record. Redistribution of stress by large earthquakes was modeled; predictions were made with and without adjustments for stress redistribution. Predictions for each segment were combined into predictions for each fault using expert opinion about the relative likelihoods of different rupture sources.

A 'time-predictable model' (stress from tectonic loading needs to reach the level at which the segment ruptured in the previous event for the segment to initiate a new event) was used to estimate the probability that an earthquake will originate on each fault segment. Estimating the state of stress before the last event requires knowing the date of the last event and the slip during the last event. Those data are available only for the 1906 earthquake on the San Andreas Fault and the 1868 earthquake on the southern segment of the Hayward Fault (*USGS*, 1999, p. 17), so the time-predictable model could not be used for many Bay Area fault segments.

The calculations also require estimating the loading of the fault over time, which in turn relies on viscoelastic models of regional geological structure. Stress drops and loading rates were modeled probabilistically (*USGS*, 1999, p. 17); the form of the probability models is not given. The loading of the San Andreas fault by the 1989 Loma Prieta earthquake and the loading of the Hayward fault by the 1906 earthquake were modeled. The probabilities estimated using the time-predictable model were converted into forecasts using expert opinion about the relative likelihoods that an event initiating on one segment will stop or will propagate to other segments. The outputs of the three types of stochastic models for each fault segment were weighted according to the opinions of a panel of fifteen experts. When results from the time-predictable model were not available, the weights on its output were in effect set to zero.

There is no straightforward interpretation of the USGS probability forecast. Many steps involve models that are largely untestable; modeling choices often seem arbitrary. Frequencies are equated with probabilities, fiducial distributions are used, outcomes are assumed to be equally likely, and subjective probabilities are used in ways that violate Bayes rule.[18]

3.1. WHAT DOES THE UNCERTAINTY ESTIMATE MEAN?

The USGS forecast is 0.7 ± 0.1, where 0.1 is an uncertainty estimate (*USGS*, 1999). The 2,000 regional models produced in stage 1 give an estimate of the long-term seismicity rate for each source (linked fault segments), and an estimate of the uncertainty in each rate. By a process we do not understand, those uncertainties were propagated through stage 2 to estimate the uncertainty of the estimated probability of a large earthquake. If this view is correct, 0.1 is a gross underestimate of the uncertainty. Many sources of error have been overlooked, some of which are listed below.

1. Errors in the fault maps and the identification of fault segments.[19]
2. Errors in geodetic measurements, in paleoseismic data, and in the viscoelastic models used to estimate fault loading and sub-surface slip from surface data.
3. Errors in the estimated fraction of stress relieved aseismically through creep in each fault segment and errors in the relative amount of slip assumed to be accommodated by each seismic source.
4. Errors in the estimated magnitudes, moments, and locations of historical earthquakes.
5. Errors in the relationships between fault area and seismic moment.
6. Errors in the models for fault loading.
7. Errors in the models for fault interactions.
8. Errors in the generic Gutenberg-Richter relationships, not only in the parameter values but also in the functional form.
9. Errors in the estimated probability of an earthquake not associated with any of the faults included in the model.
10. Errors in the form of the probability models for earthquake recurrence and in the estimated parameters of those models.

4. A view from the past

Littlewood (1953) wrote:

Mathematics (by which I shall mean pure mathematics) has no grip on the real world; if probability is to deal with the real world it must contain elements outside mathematics; the *meaning* of 'probability' must relate to the real world, and there must be one or more 'primitive' propositions about the real world, from which we can then proceed deductively (i.e. mathematically). We will suppose (as we may by lumping several primitive propositions together) that there is just one primitive proposition, the 'probability axiom,' and we will call it A for short. Although it has got to be *true*, A is by the nature of the case incapable of deductive proof, for the sufficient reason that it is about the real world ....

There are 2 schools. One, which I will call mathematical, stays inside mathematics, with results that I shall consider later. We will begin with the other school, which I will call philosophical. This attacks directly the 'real' probability problem; what are the axiom A and the meaning of 'probability' to be, and how can we justify A? It will be instructive to consider the attempt called the 'frequency theory'. It is natural to believe that if (with the natural reservations) an act like throwing a die is repeated n times the proportion of 6's will, *with certainty*, tend to a limit, p say, as $n \to \infty$. (Attempts are made to sublimate the limit into some Pickwickian sense—'limit' in inverted commas. But either you *mean* the ordinary limit, or else you have the problem of explaining how 'limit' behaves, and you are no further. You do not make an illegitimate conception legitimate by putting it into inverted commas.) If we take this proposition as 'A' we can at least settle off-hand the other problem, of the *meaning* of probability; we define its measure for the event in question to be the number p. But for the rest this A takes us nowhere. Suppose we throw 1000 times and wish to know what to expect. Is 1000 large enough for the convergence to have got under way, and how far? A does not say. We have, then, to add to it something about the rate of convergence. Now an A cannot assert a *certainty* about a particular number n of throws, such as 'the proportion of 6's will *certainly* be within $p \pm \epsilon$ for large enough n (the largeness depending on ϵ)'. It can only say 'the proportion will lie between $p \pm \epsilon$ *with at least such and such probability (depending on* ϵ *and* n_0*) whenever* $n > n_0$'. The vicious circle is apparent. We have not merely failed to *justify* a workable A; we have failed even to *state* one which would work if its truth were granted. It is generally agreed that the frequency theory won't work. But whatever the theory it is clear that the vicious circle is very deep-seated: certainty being impossible, whatever A is made to state can be stated only in terms of 'probability'.

5. Conclusions

Making sense of earthquake forecasts is difficult, in part because standard interpretations of probability are inadequate. A model-based interpretation is better, but lacks empirical justification. Furthermore, probability models are only part of the forecasting machinery. For example, the USGS San Francisco Bay Area forecast for 2000–2030 involves geological mapping, geodetic mapping, viscoelastic loading calculations, paleoseismic observations, extrapolating rules of thumb across geography and magnitude, simulation, and many appeals to expert opinion. Philosophical difficulties aside, the numerical probability values seem rather arbitrary.

Another large earthquake in the San Francisco Bay Area is inevitable, and imminent in geologic time. Probabilities are a distraction. Instead of making forecasts, the USGS could help to improve building codes and to plan the government's response to the next large earthquake. Bay Area residents should take reasonable precautions, including bracing and bolting their homes as well as securing water heaters, bookcases, and other heavy objects. They should keep first aid supplies, water, and food on hand. They should largely ignore the USGS probability forecast.

Notes

[1] See *Stigler* (1986) for history prior to 1900. Currently, the two main schools are the frequentists and the Bayesians. Frequentists, also called objectivists, define probability in terms of relative frequency. Bayesians, also called subjectivists, define probability as degree of belief. We do not discuss other theories, such as those associated with Fisher, Jeffreys, and Keynes, although we touch on Fisher's 'fiducial probabilities' in note 11.

[2] It is hard to specify precisely which conditions must be the same across trials, and, indeed, what 'the same' means. Within classical physics, for instance, if all the conditions were exactly the same, the outcome would be the same every time—which is not what we mean by randomness.

[3] A Bayesian will have a prior belief about nature. This prior is updated as the data come in, using Bayes rule: in essence, the prior is reweighted according to the likelihood of the data (*Hartigan*, 1983, pp. 29ff). A Bayesian who does not have a proper prior—that is, whose prior is not a probability distribution—or who does not use Bayes rule to update, is behaving irrationally according to the tenets of his own doctrine (*Freedman*, 1995). For example, the Jeffreys prior is generally improper, because it has infinite mass; a Bayesian using this prior is exposed to a a money-pump (*Eaton and Sudderth*, 1999, p. 849). It is often said that the data swamp the prior: the effect of the prior is not important if there are enough observations (*Hartigan*, 1983, pp. 34ff). This may be true when there are many observations and few parameters. In earthquake prediction, by contrast, there are few observations and many parameters.

[4] The number of states depends on the temperature of the gas, among other things. In the models we describe, the particles are 'non-interacting'. For example, they do not bond with each other chemically.

[5] To define the binomial coefficients, consider m things. How many ways are there to choose k out of the m? The answer is given by the binomial coefficient

$$\binom{m}{k} = \binom{m}{m-k} = \frac{m!}{k!(m-k)!}$$

for $k = 0, 1, \ldots, m$. Let n and r be positive integers. How many sequences $(j_1, j_2, \ldots, j_r)$ of nonnegative integers are there with $j_1 + j_2 + \cdots + j_r = n$? The answer is

$$\binom{n+r-1}{n}.$$

For the argument, see *Feller* (1968). To make the connection with Bose-Einstein statistics, think of $\{j_1, j_2, \ldots, j_r\}$ as a possible value of the state vector, with j_i equal to the number of particles in quantum state i.

[6] That is the number of ways of selecting n of the r states to be occupied by one particle each.

[7] In probability theory, we might think of a Maxwell-Boltzman 'gas' that consists of $n = 2$ coins. Each coin can be in either of $r = 2$ quantum states—heads or tails. In Maxwell-Boltzman statistics, the state vector has two components, one for each coin. The components tell whether the corresponding coin is heads or tails. There are

$$r^n = 2^2 = 4$$

possible values of the state vector: HH, HT, TH, and TT. These are equally likely.

To generalize this example, consider a box of r tickets, labeled $1, 2, \ldots, r$. We draw n tickets at random with replacement from the box. We can think of the n draws as the quantum states of n particles, each of which has r possible states. This is 'ticket-gas'. There are r^n possible outcomes, all equally likely, corresponding to Maxwell-Boltzman statistics. The case $r = 2$ corresponds to coin-gas; the case $r = 6$ is 'dice-gas', the standard model for rolling n dice.

Let $X = \{X_1, \ldots, X_r\}$ be the occupancy numbers for ticket-gas: in other words, X_i is the number of particles in state i. There are

$$\binom{n + r - 1}{n}$$

possible values of X. If ticket-gas were Bose-Einstein, those values would be equally likely. With Maxwell-Boltzman statistics, they are not: instead, X has a multinomial distribution. Let $j_1, j_2, \ldots j_r$ be nonnegative integers that sum to n. Then

$$P(X_1 = j_1, X_2 = j_2, \ldots, X_r = j_r) = \frac{n!}{j_1! j_2! \cdots j_r!} \times \frac{1}{r^n}.$$

The principle of insufficient reason is not sufficient for probability theory, because there is no canonical way to define the set of outcomes which are to be taken as equally likely.

[8] The most common isotope of Helium is He^4; each atom consists of two protons, two neutrons, and two electrons. He^3 lacks one of the neutrons, which radically changes the thermodynamics.

[9] There is only about one earthquake of magnitude 8+ per year globally. In the San Francisco Bay Area, unless the rate of seismicity changes, it will take on the order of a century for a large earthquake to occur, which is not a relevant time scale for evaluating predictions.

[10] The collection Σ must contain S and must be closed under complementation and countable unions. That is, Σ must satisfy the following conditions: $S \in \Sigma$; if $A \in \Sigma$ then $A^c \in \Sigma$; and if $A_1, A_2, \ldots \in \Sigma$, then $\cup_{j=1}^{\infty} A_j \in \Sigma$.

[11] Some parameters were estimated from data. The Monte Carlo procedure treats such parameters as random variables whose expected values are the estimated values, and whose variability follows a given parametric form (Gaussian). This is 'fiducial inference' (*Lehmann*, 1986, pp. 229–230), which is neither frequentist nor Bayesian. There are also several competing theories for some aspects of the models, such as the relationship between fault area and earthquake magnitude. In such cases, the Monte Carlo procedure selects one of the competing theories at random, according to a probability distribution that reflects 'expert opinion as it evolved in the study'. Because the opinions were modified after analyzing the data, these were not prior probability distributions; nor were opinions updated using Bayes rule. See note 3.

[12] About 40% of the randomly generated models were discarded for violating a constraint that the regional tectonic slip be between 36 mm/y and 43 mm/y.

[13] The standard deviations are zero—no uncertainty—in several cases where the slip is thought to be accommodated purely seismically; see Table 2 of (*USGS*, 1999). Even the non-zero standard deviations seem to be arbitrary.

[14] It seems that the study intended to treat as equally likely all $2^n - 1$ ways in which at least one of n fault segments can rupture; however, the example on p. 9 of *USGS* (1999) refers to 6 possible ways a three-segment fault can rupture, rather than $2^3 - 1 = 7$, but then adds the possibility of a 'floating earthquake', which returns the total number of possible combinations to 7. Exactly what the authors had in mind is not clear. Perhaps there is an implicit constraint: segments that rupture must be contiguous. If so, then for a three-segment fault where the segments are numbered in order from one end of the fault (segment 1) to the other (segment 3), the following six rupture scenarios would be possible: $\{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 2, 3\}$; to those, the study adds the seventh 'floating' earthquake.

[15] The relationships are all of the functional form $M = k + \log A$, where M is the moment magnitude and A is the area of the fault. There are few relevant measurements in California to constrain the relationships (only seven 'well-documented' strike-slip earthquakes with $M \geq 7$, dating back as far as 1857), and there is evidence that California seismicity does not follow the generic model (*USGS*, 1999).

[16] This probability is added at the end of the analysis, and no uncertainty is associated with this number.

[17] Stage 1 produced estimates of rates for each source; apparently, these are disaggregated in stage 2 into information about fault segments by using expert opinion about the relative likelihoods of segments rupturing separately and together.

[18] See notes 3 and 11.

[19] For example, the Mount Diablo Thrust Fault, which slips at 3 mm/y, was not recognized in 1990 but is included in the 1999 model (*USGS*, 1999, p. 8). Moreover, seismic sources might not be represented well as linked fault segments.

References

Anderson, M., J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell: 1995, 'Observation of Bose-Einstein condensation in a dilute atomic vapor'. *Science* **269**, 198–201.

Eaton, M. L. and W. D. Sudderth: 1999, 'Consistency and strong inconsistency of group-invariant predictive inferences'. *Bernoulli* **5**, 833–854.

Ellsworth, W., M. Matthews, R. Nadeau, S. Nishenko, P. Reasonberg, and R. Simpson: 1998, 'A physically-based earthquake recurrence model for estimation of long-term earthquake probabilities'. In: *Proceedings of the Second Joint Meeting of the UJNR Panel on Earthquake Research.* pp. 135–149.

Feller, W.: 1968, *An Introduction to Probability Theory and Its Applications*, Vol. I. New York: John Wiley & Sons, Inc., 3rd edition.

Freedman, D.: 1995, 'Some Issues in the Foundations of Statistics'. *Foundations of Science* **1**, 19–39.

Hartigan, J.: 1983, *Bayes Theory.* New York: Springer-Verlag.

Kolmogorov, A.: 1956, *Foundations of the Theory of Probability.* New York: Chelsea Publishing Co., 2nd edition.

Lehmann, E.: 1986, *Testing Statistical Hypotheses.* New York: John Wiley and Sons, 2nd edition.

Littlewood, J.: 1953, *A Mathematician's Miscellany.* London: Methuen & Co. Ltd.

Reif, F.: 1965, *Fundamentals of Statistical and Thermal Physics.* New York: McGraw-Hill Book Publishing Co.

Stigler, S.: 1986, *The History of Statistics: The Measurement of Uncertainty before 1900.* Cambridge, MA: Harvard University Press.

USGS: 1999, 'Working Group on California Earthquake Probabilities. Earthquake Probabilities in the San Francisco Bay Region: 2000-2030–A Summary of Findings'. Technical Report Open-File Report 99-517, USGS, Menlo Park, CA.