Earthquake predictability

B.G. Luen

Department of Statistics

October 17, 2008

So people think they can predict earthquakes?

Well, generally not the smart people.

- The long-term goal of some is prediction that's sufficiently accurate to allow evacuation ahead of time.
- Others contend such precision will never be possible.
- All agree that nothing approaching this level of accuracy is not currently available.

Predictions vs forecasts

Seismologists distinguish between:

- A *prediction* - a claim that an earthquake or earthquakes of a certain magnitude will occur (or not occur, or have some stated probability of occurring) in a stated region in a stated future time range, and
- A *forecast* - a claim or set of claims pertaining to the probability of earthquakes within a certain region and future time range.
- A forecast can, in general, be used to generate a set of predictions - or even several sets!

The short-term goal: forecasts that are more useful than sensible guessing[1].

[1] "California has lots of quakes. North Dakota has not many. I predict California will continue to have more quakes than North Dakota OMG I wuz rite I can haz grant plz"

What data do people use to predict or forecast or whatever?

We can divide forecasting schemes into two types:

- Schemes using only on the history of the process: the occurrence times/locations/magnitudes/depths/googlefritz of past earthquakes
- Schemes that, *in addition* to the history of the process, use external data: ground electrical signals caused by rock stress, cloud formations, the headaches of the guy who sends us emails whenever he has said headaches. Some of these may seem far-fetched, but how do you judge who has a point and who is nuts[2]?

Mathematostatistically, we can't say much about the second type, other than it must do better than the first type to be useful.

[2] That's a rhetorical question, Scalia.

Brad's job

The major statistical question is: how do we assess such forecasts?
Interrelated subquestions:

1. How well is it possible to predict using only the seismic history?

Brad's job

The major statistical question is: how do we assess such forecasts?
Interrelated subquestions:

1. How well is it possible to predict using only the seismic history?
2. How should we perform hypothesis tests for whether a scheme does better than sensible guessing? (What do we test it against, and how do we incorporate randomness?)

The main subsubquestion I'll consider in this talk is how to create simple statistical models for testing against (and how to assess them).
Note: So far we've only thought in detail about point processes in time, not space.

Types of predictions and forecasts

1. Periodic predictions and forecasts
2. Alarm-type predictions and forecasts
3. Moving target predictions4 and forecastss

Periodic predictions

- Issue forecasts for regular intervals in time.
- Example: At the beginning of each year, issue an estimate of the probability that one or more magnitude 5.5 or greater earthquakes occur in the SF Bay Area that year.
- Repeat for a thousand years and see how well you do.

Alarm-type predictions

The simplest alarm-type predictions just tell you whether or not there's a high risk of an earthquake now or in the near future.

- may be updated irregularly - or continuously!

Alarm-type predictions

The simplest alarm-type predictions just tell you whether or not there's a high risk of an earthquake now or in the near future.

- may be updated irregularly - or continuously!

Think of an "earthquake siren": if the siren is on, there's a high risk of an earthquake now; if the siren is off, the risk is low.
You can, of course, have a more nuanced system, with different levels of alarm (compare terror alerts), or else explicitly calculate current earthquake probabilities or rates.

Moving target forecasts

These are medium- or long-term forecasts that update, whether regularly or irregularly.
Compare to weather forecasts:

Moving target forecasts

These are medium- or long-term forecasts that update, whether regularly or irregularly.

Compare to weather forecasts:

- MONDAY: Today it was sunny, we predict it'll be sunny all week.

Moving target forecasts

These are medium- or long-term forecasts that update, whether regularly or irregularly.

Compare to weather forecasts:

- MONDAY: Today it was sunny, we predict it'll be sunny all week.
- TUESDAY: Well it rained today, we're changing our prediction to rain all week.

Moving target forecasts

These are medium- or long-term forecasts that update, whether regularly or irregularly.

Compare to weather forecasts:

- MONDAY: Today it was sunny, we predict it'll be sunny all week.
- TUESDAY: Well it rained today, we're changing our prediction to rain all week.

SO! The body of this talk will look at these three kinds of forecasts in turn.

Periodic predictions

- We like periodic predictions because they're comparatively easy to test.
- Example: At the beginning of each year, make a yes/no prediction as to whether there'll be one or more magnitude 5.5 or greater earthquakes in the SF Bay Area that year.

Periodic predictions

- We like periodic predictions because they're comparatively easy to test.
- Example: At the beginning of each year, make a yes/no prediction as to whether there'll be one or more magnitude 5.5 or greater earthquakes in the SF Bay Area that year.
- To assess: draw up a two-way table (prediction/no prediction vs quake/no quake), get 1000 years of results, and compare the results to a baseline prediction scheme.

Periodic predictions: what should we compare them to?

Say you predict earthquakes in 10% of years.
How do we construct a baseline prediction scheme to compare to?

Periodic predictions: what should we compare them to?

Say you predict earthquakes in 10% of years.
How do we construct a baseline prediction scheme to compare to?
What if we select 10% of the years at random?

Periodic predictions: what should we compare them to?

Say you predict earthquakes in 10% of years.
How do we construct a baseline prediction scheme to compare to?
What if we select 10% of the years at random?
Every seismologist knows that earthquakes are clustered in time, so a prediction scheme that accounts for this in even a stupid way will beat this random predictor in the long run.

Periodic predictions: what should we compare them to?

Say you predict earthquakes in 10% of years.
How do we construct a baseline prediction scheme to compare to?
What if we select 10% of the years at random?
Every seismologist knows that earthquakes are clustered in time, so a prediction scheme that accounts for this in even a stupid way will beat this random predictor in the long run.
So! Let's account for clustering in a stupid way!

Periodic predictions: what should we compare them to?

Say you predict earthquakes in 10% of years.
How do we construct a baseline prediction scheme to compare to?
Create a baseline scheme such that predictions are made for the 10% of years for which earthquakes (of some minimum magnitude) occurred most recently.

- The optimal minimum magnitude may be determined empirically, or, to keep things simple, just set it equal to 5.5.
- In general we don't know in advance the percentage of years for which the scheme to be tested will predict quakes. BUT, as long as the prediction algorithm is fixed in advance, it's kosher to retrospectively use the percentage of years for which quakes were predicted.

Periodic predictions: a complication!

Problem: we need like a hundred earthquake occurrences before we have decent power in a comparison between two non-useless forecasting schemes.

Periodic predictions: a complication!

Problem: we need like a hundred earthquake occurrences before we have decent power in a comparison between two non-useless forecasting schemes.

Possible solns:

1. Lower the magnitude of earthquake you're trying to predict, and reduce the period of forecasts accordingly. (Although: the best predictor at one cut-off magnitude might not work at all at a lower cut-off magnitude.)

Periodic predictions: a complication!

Problem: we need like a hundred earthquake occurrences before we have decent power in a comparison between two non-useless forecasting schemes.

Possible solns:

1. Lower the magnitude of earthquake you're trying to predict, and reduce the period of forecasts accordingly. (Although: the best predictor at one cut-off magnitude might not work at all at a lower cut-off magnitude.)
2. Make predictions for a bunch of different places simultaneously, and aggregate the results. (Although: prediction schemes specifically designed for one location may not work elsewhere. And: we still might not have much power if some locations are much more earthquake-prone than others.)

Periodic predictions: a further complication!

What if our predictions are probabilistic?

- 2009: 10% chance of a magnitude 5.5 or greater quake
- 2010: 7% chance
- 2011: 8% chance

  etc.

How do we assess such predictions?

Determinising probabilistic predictions

We can turn probabilistic predictions into 0-1 predictions quite simply:

1. Choose a cutoff probability
2. For each year, if the model probability is greater than the cut-off probability, predict an earthquake
3. If the model probability is less than the cut-off probability, predict no earthquake

But we don't want to study just one cut-off probability: we want to study *all* cut-off probabilities.

The error diagram

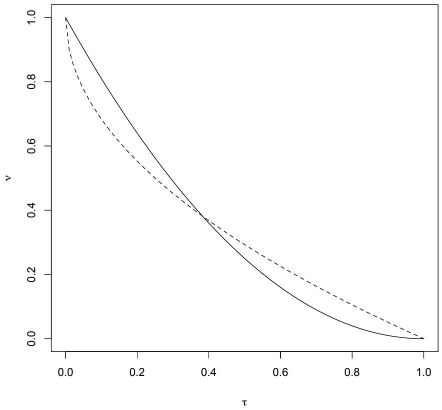

Figure: Two example error diagram curves. x-axis: τ: proportion of time covered by predictions. y-axis: ν: proportion of earthquakes not predicted.

Notes on the error diagram

- This technique doesn't explicitly account for the probabilistic nature of the predictions. There needs to be a supplementary check that the probabilities are correctly calibrated.
- It's often hard to say whether one prediction scheme performed better than another based on their error diagrams (see previous figure). We'll look at some other criteria for comparison later.

Alarm-type predictions

We can assess these much as we would periodic predictions.

- For on/off siren predictions, draw up a two-way table.
- For multi-level or probabilistic/rate predictions, draw an error diagram.

But what do baseline predictor should we test against?

Automatic alarms

- Every time an earthquake of sufficient magnitude occurs, begin an alarm.
- Fiddle the cut-off magnitude and alarm length (which may be a function of magnitude) such that the total proportion of time covered by alarms (τ) equals τ for the predictor to be tested.
- The optimal combination of cut-off magnitude and alarm length may be estimated empirically from the history of the process.
- For probabilistic predictions, repeat for all τ.

(We can also implement a "delay" parameter, so that the alarm is not declared straight away, but this isn't really useful.)

Modelling earthquake occurrence: renewal processes

Let's change gears for a moment. What's a simple way of modelling earthquake occurrence as a point process, without worrying about the size of events?

Modelling earthquake occurrence: renewal processes

Let's change gears for a moment. What's a simple way of modelling earthquake occurrence as a point process, without worrying about the size of events?

- Create a model where the time between large earthquakes is iid
- This is (one kind of) a *renewal process*

The connection between automatic alarms and renewal processes

A renewal process may be defined either in terms of a pdf for the interevent times, or a conditional intensity that's a function of the times since the last event.

- If the conditional intensity is non-increasing, then for any τ, an optimal strategy to minimise ν (based only on the history of the process) is an automatic alarm strategy where an alarm is declared immediately following an earthquake.
- It's not sufficient for the pdf to be non-increasing, eg uniform renewal times.

Simulating Bay Area earthquakes

We use a gamma renewal process with parameters $\kappa = 0.385$ and $\lambda = 0.000126$.

- meant to approximate the occurrence of magnitude 6.0 or greater earthquakes

Apply an automatic alarm strategy: after every event, declare an alarm of length 21 days (first version) or 365 days (second version).

Simulation results!

Window length	Statistic	Simulation	Theory
21 days	ν	0.9384	0.940
	τ	0.00346	0.00323
	Alarm rate	0.000162 alarms/day	0.000154
	False alarm rate	0.882	0.885
365 days	ν	0.799	0.802
	τ	0.0501	0.0479
	Alarm rate	0.000128 alarms/day	0.000131
	False alarm rate	0.664	0.660

Table: Results of semiautomatic alarm simulation for gamma renewal process with $\kappa = 0.385, \lambda = 0.000126$ and probability of alarm 0.5.

If we wanted to, we could simulate a whole error diagram.

So, why did we bother?

- Some seismologists want to know how the "intrinsic predictability" of earthquake occurrence.
- For renewal processes, it's possible to find some measures of intrinsic predictability from theory and/or simulations.

So, why did we bother?

- Some seismologists want to know how the "intrinsic predictability" of earthquake occurrence.
- For renewal processes, it's possible to find some measures of intrinsic predictability from theory and/or simulations.
- For real earthquake processes, it's not entirely clear what "intrinsic predictability" means...

Are automatic alarms best?

- Some preliminary messing around suggests that automatic alarms are about as good as we can do, in terms of predicting only based on past seismicity.
- There are physical reasons to suggest there might be some degree of periodicity to the occurrence of non-clustered earthquakes, but the empirical evidence is scant.
- We would like some assessment system sensitive enough to pick up on small improvements on automatic alarms.

Point process entropy

- Recall or learn that Shannon's entropy told you the number of bits you needed, on average, to encode a (discrete) random variable.
- This doesn't work for continuous variables, as in general you need an infinite number of bits to encode them.
- What we can define is *relative entropy*, which measures information compared to a reference measure:

$$H(\mathcal{P};\mu) = -\int \frac{d\mathcal{P}}{d\mu}(\omega)\log\frac{d\mathcal{P}}{d\mu}(\omega)\mu(d\omega) \quad (1)$$

$$= \mathbf{E}_{\mathcal{P}}\left(-\log\frac{d\mathcal{P}}{d\mu}\right) \quad (2)$$

Point process entropy

- Consider a finite point process on A with indistinguishable points.
- Let R be the number of points in A, and $p_r = P(R = r)$.
- The likelihood of a realisation of the process can be written as $p_r r! s_r(x_1, \ldots, x_r; A)$, where s_r is a symmetric probability density.

Point process entropy

- Consider a finite point process on A with indistinguishable points.
- Let R be the number of points in A, and $p_r = P(R = r)$.
- The likelihood of a realisation of the process can be written as $p_r r! s_r(x_1, \ldots, x_r; A)$, where s_r is a symmetric probability density.

Neglecting the outcome of zero events, the entropy is:

$$\begin{aligned} H(A) &= -\mathbf{E} \log(p_R R! s_R(\mathbf{X})) = -\mathbf{E}[\mathbf{E}[\log(p_r r! s_r(\mathbf{X}))|R = r]] \\ &= -\sum_r p_r \log(r! p_r) - \sum_r \left(p_r \int s_r \log s_r d\mathbf{x} \right) \end{aligned}$$

The information score

One entropy-based measure of predictability employed by seismologists is the *information score*.
In discrete time, it's

$$I_i = p_i \log \frac{p_i}{\pi_i} + (1 - p_i) \log \frac{1 - p_i}{1 - \pi_i}. \tag{3}$$

This is just the Kullbach-Leibler divergence of the baseline model π_i at the prediction model p_i, for a single cell i.
Major caveat: it only gives useful results if the prediction model is "close" to "true".

The information score

The information score for a continuous-time point process model is

$$\hat{I} = \frac{1}{N}\sum \log \frac{\lambda_i}{\bar{\lambda}} \tag{4}$$

where the average is taken over the number of earthquakes.
Here λ_i is the predicted rate of earthquakes at the time the ith quake occurs. λ is the baseline rate, usually taken to be constant (implicitly assuming a Poisson process).
If the model predicts earthquakes at the same average rate as they actually occur, this score is bounded above by the relative entropy (rescaled by a factor of $1/\lambda$).
If the model predicts earthquakes at a incorrect average rate, it seems we can just append a correction term that integrates the difference in rates, but I haven't tested this.

Moving targets

- No real predictor updates continuously: there's always some lag time.
- Furthermore, some lag time is obviously necessary for a forecast to be useful. Saying “Hey everybody, there's going to be an earthquake two seconds from ¡¡¡BLAM!!!” is not a useful prediction[3].
- Furtherfurthermore, most medium- and long-term forecasting schemes change their mind: newly issued predictions supersede old ones.

[3] Although Richard Allen, who's on my committee and is trying to do more or less this, argues it may be useful in practice.

Moving targets

So how best to assess moving target forecasts?

- In weather forecasting, forecasts at different lead times are usually assessed separately: so you assess your one day ahead forecasts, your two day ahead forecasts, etc.
- There's a multiple testing issue that (in my limited reading) hasn't been addressed. Obviously the forecasts aren't independent, but can we pretend they are?

Moving targets

Is there a way to measure how well a moving target scheme performs without specifying an arbitrary lead time?

- I've played around a tiny bit with cost-benefit analyses, but it's really hard to get convincing ballpark estimates of costs and benefits.
- At this stage, should we just take the set of most recent predictions as the model's best set of predictions? Do you lose much by doing this?

Future work

How do we account for earthquake magnitude?

- Marked point processes would work. What's the simplest extension of a renewal process that gives sensible marks?
- The ETAS model is a well-liked seismologist-developed "cascade"-type model that accounts for magnitude, but it's a bit complicated.

Future work

The spatial stuff.

- We want to determine what spatial extent automatic alarms should take.
- If this is fixed, then it's an easy empirical problem.
- If this is a function of magnitude, then it's a harder empirical problem.
- If we want the shape of the alarm to depend on geography/past seismicity, it gets messy.

Future work

Hey, you remember how in my 2005 student talk, I said what I really need to do was take the plunge and apply our ideas to a real-life predictor?

Future work

Hey, you remember how in my 2005 student talk, I said what I really need to do was take the plunge and apply our ideas to a real-life predictor?

- Well, I still need to do that.

Fin.