MS&E 347 Term Paper on Pricing Earthquake Reinsurance

Nathan Brown
Ophir Gottlieb
Ari Officer
Anca Vacarescu
V Vishal

June 14, 2007

Abstract

The pricing of tranches on earthquake claims offers the dual challenge of pricing derivatives in an incomplete market with non-tradable underlying. Here we adopt a utility based approach to empirically account for the risk aversion of investors in the market to price tranches. Earthquake occurences are modelled statistically with variants of the Hawkes Process. Monte Carlo simulations are then used to obtain the price of tranches for different market parameters.

1 Introduction

Tranches are instruments sold by insurance companies to reduce their exposure to loss. By collecting together a large pool of retail insurance contracts and restructuring the portfolio into tranches, they offer investors the opportunity to trade in instruments with various risk/return profiles, while reducing their own exposure to losses.

When the insurance is offered on tradable assets (such as corporate bonds), the arbitrage-free price of the tranches can be obtained from traditional risk-neutral pricing arguments. However, in earthquake insurance, where the underlying is a non-tradable phenomenon, there is no natural martingale measure that can be used to price derivatives. In this paper, we price a tranche on California earthquake claims using the approach of Davis[1]. For completeness, we review the principal arguments of his method. Consider an investor with a concave utility function $U(x)$ and initial endowment η, holding an optimal portfolio of traded assets that maximizes the expected utility. Let the value process of the portfolio be $H_t^*(\eta)$. Given some fixed time horizon T, we say that the price p of an asset with value B_t is 'fair' if the investor is neutral to the diversion of a small portion of his portfolio into the contingent claim. i.e. the fair price is the value of p_t at which

$$\frac{\partial}{\partial \delta} \mathrm{E}\left[U(H_T^*(\eta - \delta) + \frac{\delta}{p_t} B_T) \;\Big|\; \mathcal{F}_t \right] \Bigg|_{\delta=0} = 0$$

This leads to the following pricing formula:

$$p_t = \frac{\mathrm{E}\left[U'(H_T^*(\eta)) B_T \;\Big|\; \mathcal{F}_t \right]}{\mathrm{E}\left[U'(H_T^*(\eta)) H_T^{*\prime}(\eta) \;\Big|\; \mathcal{F}_t \right]} \tag{1}$$

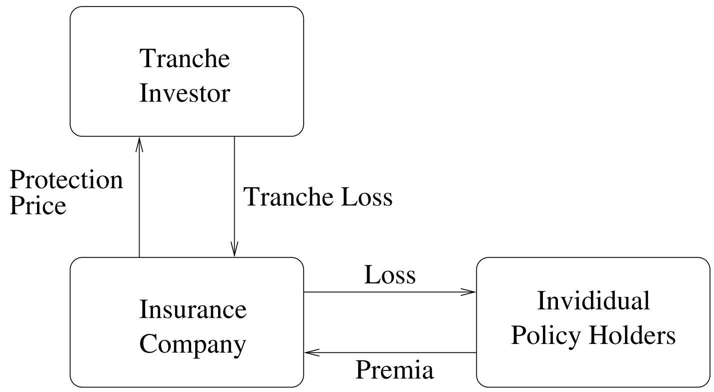


Figure 1: Relationship between parties involved in earthquake insurance: The invidual policy holders pay a premium stream to the insurance company and the insurance company pays damages to the policy holder. The insurance company aggregates the individual policies and restructures the aggregate losses into tranches that are of interest to different types of investors. We wish to determine the fair protection price for a given tranche.

Thus to price a tranche on earthquake claims, we need

i the distribution of the tranche loss process B_t

ii a suitable concave utility function U

iii a suitable optimal reference portfolio H_t^*

2 Modeling Earthquake Occurence

We examined earthquake occurrences from the perspective of jump processes. Data available to us included all recorded seismic activity in the state of California over the period 1975-2005. Shown in fig. 2 is a map of california showing locations of occurrence in the period 1991-2001[2]. There is significant spatial variation in frequency and intensity and in the following development, we explore the possibility of improving fits by parametrising the statistical models by spatial location. Shown in fig. 3 is the earthquake occurrence data for a location east of Los Angeles. Below, we consider several classes of jump processes as models for the data.

Homogeneous Poisson Process

The simplest model is a homogeneous Poisson process with independently distributed magnitudes. To establish notation, we define the process here: The homogeneous Poisson process with intensity λ is a counting process N_t that has independent and identically distributed increments in equal non-overlapping intervals of time with the increment in a given time interval $(s,t]$ given by the Poisson distribution:

$$P_s[N_t - N_s = k] = \frac{1}{k!}(\lambda(t-s))^k e^{-\lambda(t-s)} \qquad k = 0, 1, 2, ...$$

Each sample path $N_t(\omega)$ is piecewise constant, nondecreasing, right continuous with $N_0(\omega) = 0$ and with jumps of size one.

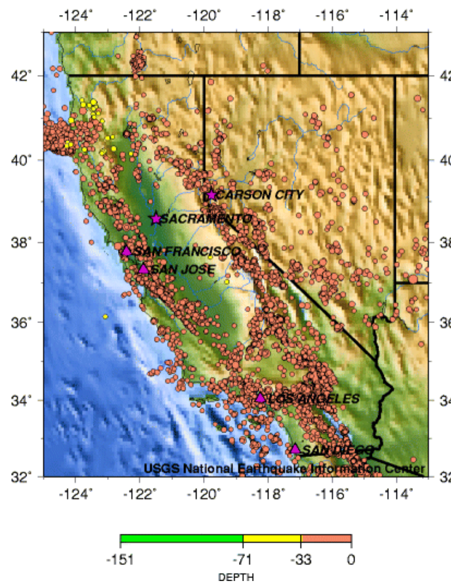


Figure 2: Historical map of seismic activity in California (1990-2000)

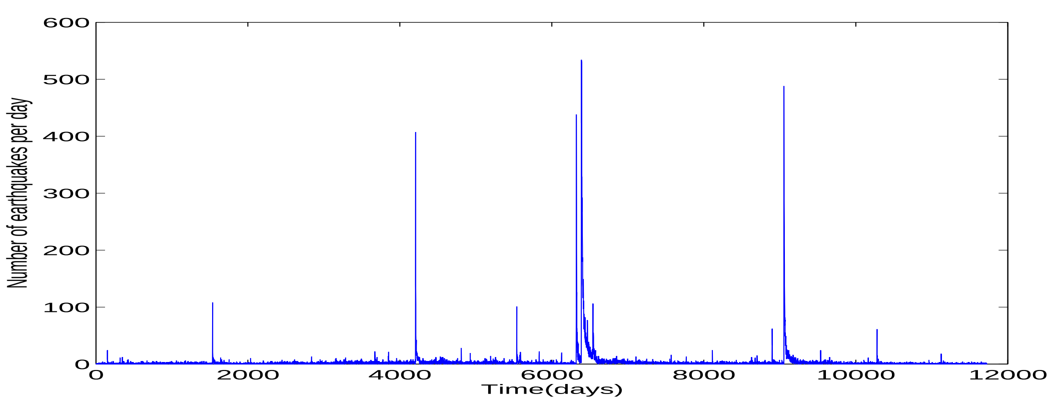
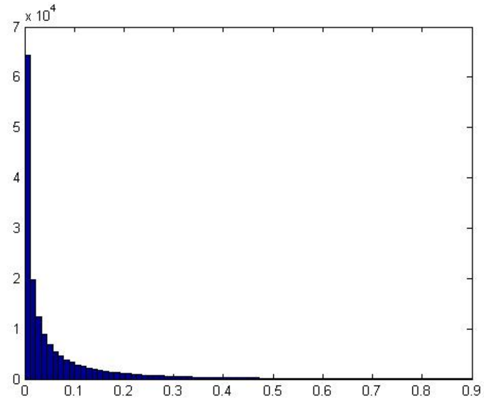
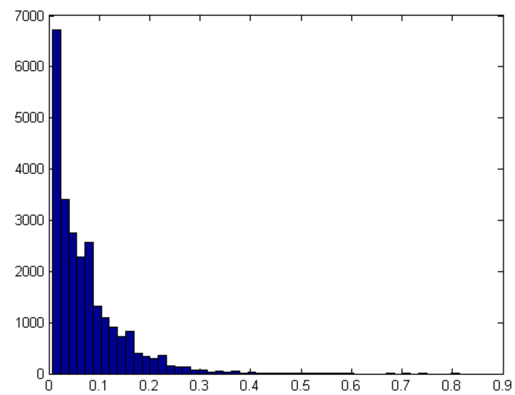


Figure 3: Frequency of earthquake occurence over 30 years for location east of Los Angeles - Can we make the axes labels larger?

The interarrival times of the observed time series are shown in fig. 4(a). Compare this to the interarrival times for the homogeneous Poisson process simulated from the MLE estimate λ shown in fig. 4(b). Visually the comparison is compelling and suggests that a homogenous Poisson process might be an adequate model. To make the comparison more concrete,fig. 5 is a quantile-quantile(QQ) plot of the interarrival times. We see that the homogenous Poisson process fits the smaller interarrival times best while the predicted longer interarrival times are much longer than those observed.



(a)



(b)

Figure 4: Histogram of interarrival times for observed data (left) and simulated data from Poisson process(right).

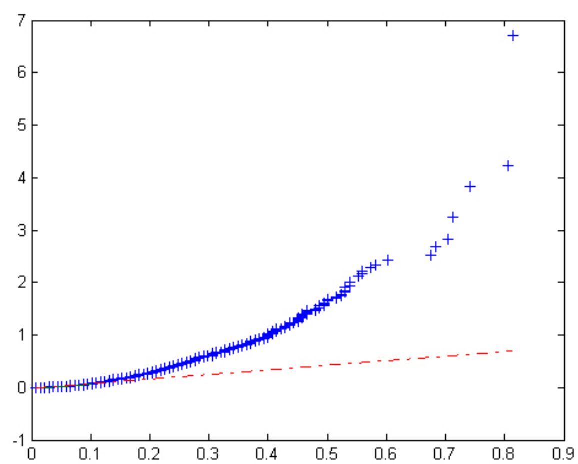


Figure 5: QQ plot of interarrival times for observed data (y-axis) vs. simulated data(x-axis)

Jump Processes with Stochastic Intensity

A more general model for the counting process N_t of earthquake occurrences can be developed by allowing the intensity λ to vary stochastically as a function of some jump diffusion X. In complete generality, we consider $\lambda = \lambda(X_t, t)$, where X is a strong solution to the stochastic differential equation

$$dX_t = \mu(X_t, t)\, dt + \sum_{j=1}^{p} \sigma^j(X_t, t)\, dW_t^j + \sum_{i=1}^{m} \zeta^i\, dZ_t^i,$$

Here the W^j are $\mathbb{R}^d$ standard Brownian motions with any given correlation matrix, the ζ^i are d-dimensional diagonal matrices, and each Z^i is a temporally consistent $\mathbb{R}^d$-valued point process. In other words, the component processes of each vector Z^i share event times.

The point process characteristics of N_t are determined in this setting from the interdependence structure established between N and its driving factor X. If X is independent of N, then so is λ, and feedback effects are ruled out. In this case, N is called a doubly stochastic point process. If X, and hence λ depend on N, then the process is self-affecting and has the ability to capture feedback. This is the case for example when N is temporarily consistent with one of the component processes of one of the jump terms Z^i of X.

To formalize the discussion above, let $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ and $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ denote the right-continuous and completed version of the filtration generated by X, and N respectively. Here, $(*_t)_{t \geq 0}$ are nondecreasing, time-indexed collection of sub σ-algebras of the corresponding σ-algebra, encoding the information available up to time t.

In the doubly stochastic setting, we know that the dynamics of N have no effect on the process X; but at any time s<t, conditional on $\mathcal{F}_t$, the process N is a Poison process with rate λ_s:

$$P_s[N_t - N_s = k \mid \mathcal{G}_s \vee \mathcal{F}_t] = \frac{1}{k!}\Big(\int_s^t \lambda_s^i\, ds\Big)^k e^{(\int_t^{t+u} \lambda_s^i\, ds)} \qquad k = 0, 1, 2, ...$$

In a doubly stochastic credit model, X typically consists of historical default information, credit sensitive security prices, macro-economic variables and other information that is more-or-less easy to observe. However, in the case of earthquakes, driving factors would likely comprise of complex dynamical variables related to the physics of the earth's core and crust. Since such variables are not readily observable and the physics is complex, we might instead consider a statistical black box approach. For example, we might assume a mean reverting process for the intensity with parameters to be determined by fitting to the observed seismicity data. Although the idea seems sound, it was not possible within the scope of this term project to pursue this direction and we leave doubly stochastic processes as a possibility for future work.

Hawkes Process

It is not unreasonable to expect that the occurence of an earthquake will affect the likelihood of future earthquakes. For example, earthquakes typically have aftershocks, which represents a temporary increase in the intensity at each arrival time. Consequently we consider the class of self-affecting jump processes mentioned earlier. The driving process filtration $\mathbb{F} \subseteq \mathbb{G}$ is the complete, right continuous filtration generated by N itself. The proces intensity λ is adapted to $\mathbb{F}$: only information generated by the defaults up to time t is used to model the intensity λ_t.

The simplest self-affecting stochastic process is the standard birth process defined as a point process whose intensity is an increasing function of the counting process itself. The intensity for an affine birth process is of the form:

$$\lambda_t = \lambda_\infty + \delta N_t$$

where $\lambda_\infty > 0$ is the intensity of the first event and $\delta \geq 0$ is the feedback sensitivity to an event. Larger δ implies greater feedback (intensity increases by δ for each jump). We see that the case when $\delta = 0$ takes us back to the homogeneous Poisson process we examined above (where $\lambda_\infty =$ constant). Extending this birth process yields the Hawkes process which can be written generally as:

$$\lambda_t = \lambda_\infty + \int_0^t d(t-s)dN_s$$

where λ_∞ is some positive constant representing the intensity to the first earthquake. There are many choices for the deterministic response function $d(t)$; for our purposes we chose $d(t) = \alpha e^{-\beta t}$ which offers easy interpretability. Specifically the intensity grows by α instantaneously after a seismic shock and the impact of that event decreases exponentially at a rate of β. Finally we can write the Hawkes intensity dynamics as an SDE, noting that the special case when $\beta = 0$ brings us back to the standard birth process (see section 6 for details):

$$d\lambda_t = \beta(\lambda_\infty - \lambda_t)dt + \alpha dN_t$$

$$\lambda_0 = \lambda_\infty$$

In the Hawkes process, the only randomness in the intensity comes from the jump process itself. The feedback property provides us the ability to model earthquake clusters. In this model, macro-geophysical shocks arrive according to a Poisson process with intensity λ_∞, while an earthquake can initiate a cluster of new earhquakes. These clustered events are themselves specified by an independent inhomogeneous Poisson process with intensity $d(t)$. The Hawkes process N_t is the overlapping of the original λ_∞ intensity Poisson process with the triggered inhomogeneous Poisson process. Figures 6 and 7 present one sample path of a Hawkes process and a homogeneous Poisson Process for comparison. We can observe the clustering property in the Hawkes process. Figures 8 and 9 illustrate the histograms of the simulated Hawkes process for earthquake inter-arrival times for one location and the entire state respectively. An examination of the histogram of actual inter-arrival times from above shows a much closer fit to the Hawkes simulated arrival times than that of the homogeneous Poisson process.

On the use of a Hawkes Process to model earthquakes

We had the opportunity to discuss the Hawkes process as a model for earthquake occurrences with Dr. Mary Lou Zoback, senior research scientist of the U.S. Geological Surveys Earthquake Hazards Team of Menlo Park, California. According to Dr. Zoback, the Hawkes process even in a more general form falls short of describing the actual complexity of earthquake interactions. On a positive note, the exponential decay used in this particular Hawkes process corresponds to the accepted model for energy loss after a quake, which results in aftershocks along the same fault line. However, the further damage from such aftershocks is generally negligible. The USGS discounts all aftershock data in its own modeling. After these aftershocks, the probability of another earthquake occurring along the same fault actually decreases tremendously, as the pressure that caused the earthquake has been released.

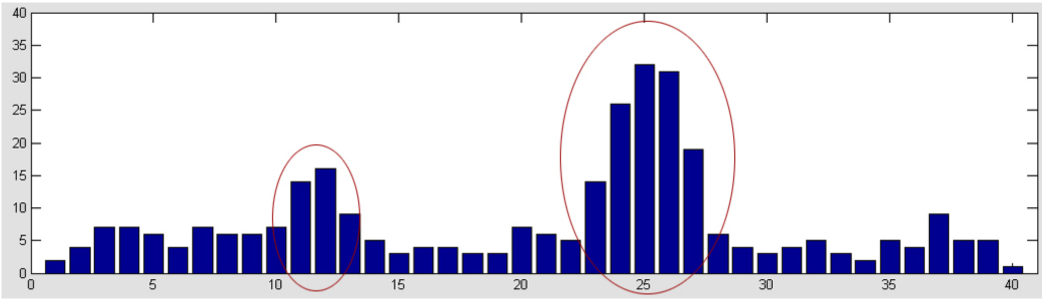


Figure 6: One sample path of a Hawkes Process

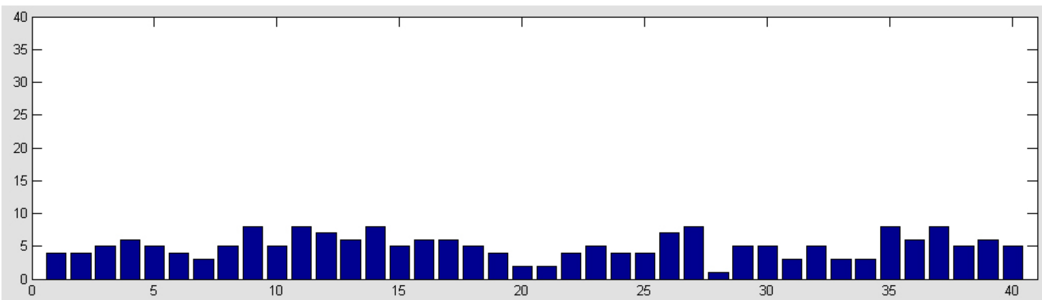


Figure 7: One sample path of a Homogeneous Poisson Process

Between faults, self-excitement of the earthquakes is even less clear. Sometimes, when an earthquake releases the pressure built along its fault, it creates new pressure on all of the surrounding faults; other times, it can actually release the pressure. Thus, the dependence among earthquake occurrences on different faults is not geologically sound. We have two historical examples illustrating these extreme cases: The 19th century as a whole was much more active than the 20th century for earthquakes in the San Francisco Bay area. The famous 1906 San Francisco earthquake was so enormous that it took much of the pressure off of the faults around the Bay area, resulting in no major quakes in the area until 1989. On the other hand, seismologists believe that the Mojave Earthquake of 1992, which occurred July 11, was caused by the Landers Earthquake, which occurred June 28. These both occurred in an area in which geologists expected only one earthquake in a thousand years.

Considering the only certain effects of an earthquake lie in its aftershocks, it makes sense that the Hawkes process predicts a greater dependency among earthquakes over longer periods of time than is expressed empirically. However, for pricing purposes, the shortcomings of the Hawkes process in measuring the intricate dependence are not necessarily detrimental. Since effects among faults are random events, taking the maximum likelihood over the data takes into account that randomness in some sense. Thus, because we cannot produce a model that accounts for the complex geological interactions, we stick with the Hawkes process as a model for earthquake occurrences but provide this explanation as a caveat.

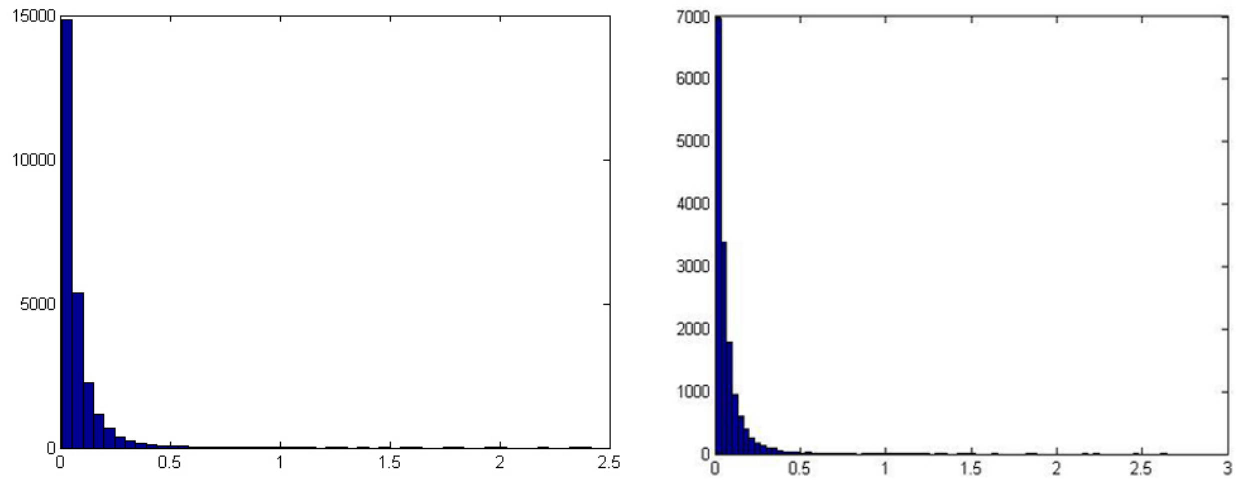


Figure 8: Histogram of interarrival times for Hawkes process (right) compared to histogram for observed data.

3 Modeling the losses due to earthquakes

We imposed an assumption on the loss characteristics in order to transition from earthquake magnitudes to the losses; specifically, we felt an S-shaped curve fit well with empirical data. This assumption was formed by using empirical data provided by US Geological Survey documents summarized on various source websites[1]. An S-shape curve imposes asymptotic behavior on both the low and high end of the magnitude spectrum while also allowing for exponential loss growth in the middle of the spectrum which the empirical data suggest.

Noting that the largest earthquake in magnitude ever recored was 9.5 (Chile 1960), the largest and most devastating earthquake in California (1906 San Francisco) history measured 7.8 and the Loma Prieta earthquake (1989 Northern California) measured 7.0, suggests asymptotic near 7.0 in magnitude on the high end. Further, earthquakes below 3.0 rarely cause damage and those below 2.0 are rarely felt, suggesting asymptotic behavior near 3.0 on the low end.

Since the loss can occur at various times, for simplicity, we shift all losses to time T by multiplying with the factor e^{rT}.

4 Pricing

In order to obtain prices, the expectations in eq. 1 need to be obtained using a Monte Carlo procedure. We adopted the thinning approach of Lewis and Shedler[3]. A brief overview of the method and some other techniques we tried are described in the appendix.

Once a model for the earthquake process has been adopted, we need to choose a reference portfolio H_T^* and a utility function $U(x)$ to obtain prices according to eq. 1. From the point of view of the reinsurer, we choose the optimal portfolio to be the market portfolio, which is completely independent of the earthquake loss process. The underlying assumption here is that there are no readily traded assets available to the reinsurer that are highly correlated with the loss process (other than the tranches we are trying to price). Consequently, the risk aversion of the reinsurer is

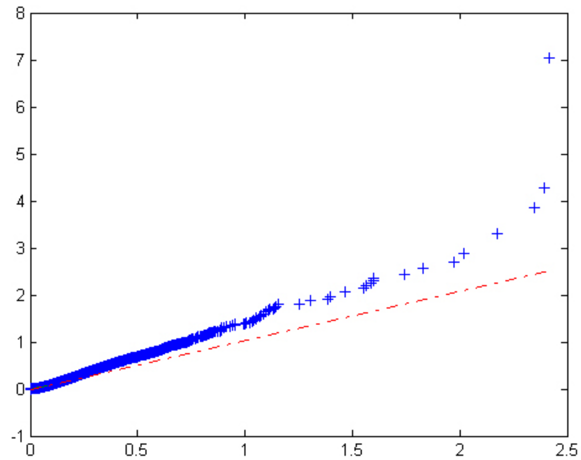


Figure 9: QQ plot of Observed interrarrival times (y-axis) vs. simulated ones from Hawkes process(x-axis)

assumed to be entirely captured by their investment preferences in the regular market. One might be able to get more realistic prices by identifying traded assets that are strongly correlated with the loss process. (Positive correlations would lead to higher values and negative correlations would lead to lower values in eq. 1. Ultimately of course, the true price is not unique in the incomplete market.)

If we assume that the reinsurer has a logarithmic utility and that the optimal portfolio can be represented by a single geometric brownian motion with growth rate μ, then formula 1 reduces to

$$p_t = e^{-\mu T} E[f(L_T; K_l, K_u)|\mathcal{F}_t] \tag{2}$$

where $f(x; K_l, K_u) = (x - K_l)^+ - (x - K_u)^+$ is the tranche loss with attachment point K_l and detachment point K_u. Compare this to the formula for risk neutral pricing:

$$p = e^{-rT} E^Q[f(L_T; K_l, K_u)]$$

price structure for the tranches with a time horizon of 1 year (in the same units as the loss curve in fig. 10). One interesting point to note here is that equation 2 is only valid when f is non-replicable. If it is possible to replicate the claim, then the optimal portfolio consists of the an additional amount invested in the μ portfolio, the inital cash outlay for which is obtained from short sale of the replicating portfolio (see Davis[1]).

We can also try to find the fair price from the point of view of the insurance company. Suppose for simplicity that the insurance company does not engange in any business other than earthquake insurance. Then the wealth dynamics of the company could be written as

$$dH_t = \phi dt + rH_t dt - dL_t$$

$$H_0 = \eta$$

where r is the risk free interest rate, η is the initial cash holdings, ϕ is a continuous (aggregated) premium rate from policy holders and L_t is the loss process. Optimal values of ϕ and η might be

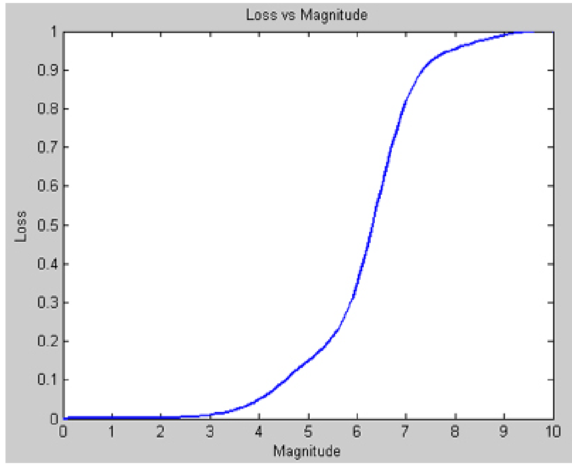


Figure 10: Magnitude to Loss conversion

determined from the market and government regulations for example. With the convention that losses are appropriately shifted to T, we can write the wealth at time T as

$$H_T = \frac{\phi}{r}(e^{rT} - 1) + \eta e^{rT} - L_T$$

In the logarithmic utility setting, the price of the tranche from the company's perspective is

$$p = e^{-rT} \frac{\mathrm{E}[\frac{L_T}{H_T}|\mathcal{F}_t]}{\mathrm{E}[\frac{1}{H_T}|\mathcal{F}_t]}$$

The use of the logarithmic utility captures the insurance company's aversion to bankruptcy. In cases where the final wealth is close to zero, the price the company is willing to pay for protection increases inversely. (Note we have not modelled any constraints on bankruptcy at times $t < T$, only at T.)

Now, in order for the tranches to have a market, the fair price according to the tranche investor should be less than or equal to the fair price according to the company. This can be used as additional constraints in determining η and ϕ by the insurance company.

As a numerical demonstration of the concepts involved, we chose the following tranche structure and market parameters: Since the minimum loss is much greater than zero, we assume a 'deductible' for the insurance company of $D = \mathrm{E}[L_T] - 3\sigma(L_T)$ where $\sigma(X) = \sqrt{\mathrm{Var}[X]}$. i.e. the insurance company does not securitize losses up to D. The nominal on the remaining loss above the deductible is assumed to be $\mathrm{E}[L_T - D] + 3\sigma(L_T)$. The growth rate of an optimal portfolio available in the market is assumed to be $\mu = 15\%$ while the risk free interest rate was set at 5%. The initial cash outlay for the insurance company is assumed to be $\eta = 0.3\,\mathrm{E}[L_T]$ and the premium rate is assumed to be such that $\phi T = \mathrm{E}[L_T] + 3\sigma(L_T)$. With these settings, we obtain the following prices for the tranches from the perspective of each party (prices are fraction of nominal):

Tranche	Insurance Company would pay	Tranche Investor demands
0 - 3%	0.0234	0.0142
3 - 7%	0.0311	0.0189
7 - 10%	0.0233	0.0142
10 - 15%	0.0388	0.0235
15 - 30%	0.1127	0.0679
$> 30\%$	0.1734	0.0974

For these particular values, there is a market for each tranche since the investors fair price is below the price the insurance company would be willing to pay. Interestingly if we assume that the insurance company also has access to the optimal portfolio of the investor at rate μ, then the prices according to the insurance company turn out to be as listed below:

Tranche	Insurance Company would pay	Tranche Investor demands
0 - 3%	0.0142	0.0142
3 - 7%	0.0189	0.0189
7 - 10%	0.0141	0.0142
10 - 15%	0.0235	0.0235
15 - 30%	0.0681	0.0679
$> 30\%$	0.0971	0.0974

Although not immediately obvious why, this choice of wealth process results in nearly identical prices from the point of view of the investor and the insurance company! Economically, this leads to a consistent interpretation - given identical investment opportunities and utilities, the investor and insurance company have nearly identical risk aversion and hence price the tranches almost identically.

5 Acknowledgements

We would like to thank Prof. Kay Giesecke and Xiaowei Ding for their extremely helpful suggestions and guidance in this work.

6 Appendix

The Hawkes Process as an Affine Jump Process

We apply integration by parts for Stieltjes integrals to the definition of λ_t above and obtain the equivalent model formulation

$$d\lambda_t = \beta(\lambda_\infty - \lambda_t)\, dt + \alpha\, dN_t$$

Explicitly,

$$\begin{aligned}
\int_0^t \alpha e^{-\beta(t-s)}\, dN_S &= \alpha N_T - \alpha e^{-\beta t} N_0 - \int_0^t \alpha\beta e^{-\beta(t-s)} N_s\, ds \\
&= \alpha N_t - \beta \left(\int_0^t \alpha e^{-\beta(t-s)} N_s\, ds \right)
\end{aligned}$$

By definition,

$$d\lambda_t = \alpha\, dN_t - \beta \frac{\partial}{\partial t} \left(\int_0^t \alpha e^{-\beta(t-s)} N_s\, ds \right) dt.$$

However, we know by Duhamel's principle that $\int_0^t \alpha e^{-\beta(t-s)} N_s\, ds$ is a solution to

$$\frac{\partial u}{\partial t} = -\beta u + \alpha N_t, \quad u(0) = 0$$

Thus,

$$\frac{\partial}{\partial t} \left(\int_0^t \alpha e^{-\beta(t-s)} N_s\, ds \right) = -\beta \left(\int_0^t \alpha e^{-\beta(t-s)} N_s\, ds \right) + \alpha N_t = \lambda_t - \lambda_\infty,$$

giving the expression below.

$$\begin{aligned}
d\lambda_t &= \beta(\lambda_\infty - \lambda_t)dt + \alpha dN_t \\
\lambda_0 &= \lambda_\infty
\end{aligned}$$

Model fitting

For the Hawkes process with exponentially decaying feedback the intensity is given by

$$\lambda(t; \nu, \alpha, \beta) = \nu + \int_{-\infty}^{t-} \alpha M(u) e^{-\beta(t-u)} dN(u)$$

where $N(t)$ is the counting process for earthquake occurence and $M(t)$ is the magnitude process.

To obtain parameter estimates from the data, we follow a maximum likelihood approach. As described in Ogata[ref], the (approximate) log-likelihood function for a stationary point process observed up to time T is given by

$$L(\nu, \alpha, \beta) = -\int_0^T \lambda(t, \omega; \nu, \alpha, \beta) dt + \int_0^T \log \lambda(t, \omega; \nu, \alpha, \beta) dN(t)$$

Let t_i be the arrival time of the i^{th} earthquake. Let $t_0 = 0$ and let $T = t_N$ be the arrival time of the last observed earthquake. Then for the Hawkes process above, we have:

$$\lambda(t; \nu, \alpha, \beta) = \nu + \sum_{i=1}^{N_{t-}} \alpha M(t_i) e^{-\beta(t-t_i)}$$

$$\lambda(t_j;\nu,\alpha,\beta) = \nu + \sum_{i=1}^{j-1} \alpha M(t_i) e^{-\beta(t_j - t_i)}$$

and the log likelihood function is

$$\begin{aligned}
L(\nu,\alpha,\beta) &= -\int_0^T \left(\nu + \sum_{i=1}^{N_{t-}} \alpha M(t_i) e^{-\beta(t-t_i)} dt \right) + \sum_{i=1}^{N} \log \lambda(t_i) \\
&= -\nu T - \sum_{j=1}^{N} \int_{t_{j-1}}^{t_j} \sum_{i=1}^{N_{t-}} \alpha M(t_i) e^{-\beta(t-t_i)} dt + \sum_{i=1}^{N} \log \lambda(t_i)
\end{aligned}$$

Since $N_{t-} = N_{t_{j-1}} = j-1$ for $t \in (t_{j-1}, t_j]$, we can integrate inside the inner sum:

$$\begin{aligned}
&= -\nu T - \sum_{j=1}^{N} \sum_{i=1}^{j-1} \alpha M(t_i) \left[\frac{-1}{\beta} e^{-\beta(t-t_i)} \right]_{t_{j-1}}^{t_j} + \sum_{i=1}^{N} \log \lambda(t_i) \\
&= -\nu T - \sum_{j=1}^{N} \sum_{i=1}^{j-1} \frac{\alpha}{\beta} M(t_i) \left[e^{-\beta(t_{j-1}-t_i)} - e^{-\beta(t_j - t_i)} \right] + \sum_{i=1}^{N} \log \lambda(t_i)
\end{aligned}$$

Swapping the order of summation, we get

$$= -\nu T - \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{\alpha}{\beta} M(t_i) \left[e^{-\beta(t_{j-1}-t_i)} - e^{-\beta(t_j - t_i)} \right] + \sum_{i=1}^{N} \log \lambda(t_i)$$

The inner sum is telecscoping:

$$L(\nu,\alpha,\beta) = -\nu T - \frac{\alpha}{\beta} \sum_{i=1}^{N} M(t_i)[1 - e^{-\beta(T-t_i)}] + \sum_{i=1}^{N} \log \lambda(t_i)$$

Maximizing L gives the MLE estimates for the parameters ν, α, β. λ can be evaluated using the following recursion:

$$\lambda(t_{j+1}) = \nu + (\lambda(t_j) - \nu + \alpha M(t_j)) e^{-\beta(t_{j+1} - t_j)}$$

Directly maximizing the log likelihood function using numerical techniques runs into some numerical issues, mainly because of the logarithmic term which can cause the numerical gradient to blow up. So we consider solving the gradient equations numerically instead:

Let $C = \sum_1^N M_i[1 - e^{-\beta(T-t_i)}]$. Then $L = -\nu T - \frac{\alpha}{\beta} C + \sum_1^N \log \lambda_i$. Some useful identities:

$$\begin{aligned}
\frac{\partial \lambda_i}{\partial \nu} &= 1 \\
\frac{\partial \lambda_i}{\partial \alpha} &= \frac{\lambda_i - \nu}{\alpha} \\
\frac{\partial \lambda_i}{\partial \beta} &= -\alpha \sum_{j=1}^{i-1} M(t_j) e^{-\beta(t_i - t_j)} (t_i - t_j)
\end{aligned}$$

Now we can get the derivatives of the log likelihood function:

$$
\begin{aligned}
\frac{\partial L}{\partial \nu} &= -T + \sum_{1}^{N} \frac{1}{\lambda_i} \\
\frac{\partial L}{\partial \alpha} &= -\frac{1}{\beta} C + \frac{1}{\alpha} \sum_{i=1}^{N} \left(1 - \frac{\nu}{\lambda_i}\right) \\
\frac{\partial L}{\partial \beta} &= \frac{\alpha}{\beta^2} C - \frac{\alpha}{\beta} \sum_{i=1}^{N} M(t_i)(T - t_i) e^{-\beta(T - t_i)} - \alpha \sum_{i=1}^{N} \frac{1}{\lambda_i} \sum_{j=1}^{i-1} M(t_j) e^{-\beta(t_i - t_j)} (t_i - t_j)
\end{aligned}
$$

Let $\epsilon_{ij} = 1$ if $j < i$ and zero otherwise. Then we can write

$$
\begin{aligned}
\frac{\partial L}{\partial \beta} &= \frac{\alpha}{\beta^2} C - \frac{\alpha}{\beta} \sum_{i=1}^{N} M(t_i)(T - t_i) e^{-\beta(T - t_i)} - \alpha \sum_{i=1}^{N} \frac{1}{\lambda_i} \sum_{j=1}^{N} \epsilon_{ij} M(t_j) e^{-\beta(t_i - t_j)} (t_i - t_j) \\
\frac{\partial L}{\partial \beta} &= \frac{\alpha}{\beta^2} C - \frac{\alpha}{\beta} \sum_{i=1}^{N} M(t_i)(T - t_i) e^{-\beta(T - t_i)} - \alpha \sum_{i=1}^{N} \frac{1}{\lambda_i} \sum_{j=1}^{N} \epsilon_{ij} M(t_j) e^{-\beta(t_i - t_j)} (t_i - t_j)
\end{aligned}
$$

Simulation

We mainly used the thinning procedure of Lewis et al.[3] for simulating the Hawkes process. This was found to be quite efficient for a large number of simulations. However, using this method requires being able to bound λ in each interarrival period. So we also explored a simpler Euler discretized version to simulate the process λ and the jump process together. For a Hawkes process with exponential coupling, we calculate the increments using

$$
\begin{aligned}
dN &= \begin{cases} 1 & \text{with probability} \quad \lambda_t dt \\ 0 & \text{with probability} \quad 1 - \lambda_t dt \end{cases} \\
d\lambda &= -\beta \lambda dt + \alpha dN
\end{aligned}
$$

and compute trajectories with

$$
\begin{aligned}
t_{i+1} &= t_i + dt \\
N_{i+1} &= N_i + dN \\
\lambda_{i+1} &= \lambda_i + d\lambda
\end{aligned}
$$

This makes it easy to adapt the code to different dynamics for λ, with the caveat that the time step may need to be fairly small to ensure accurate integration of the the λ SDE. To speed things up a little bit, we employed an adaptive time step method that uses as large a dt as possibly, while maintaining an upper bound on the product λdt to ensure that there was no significant probability

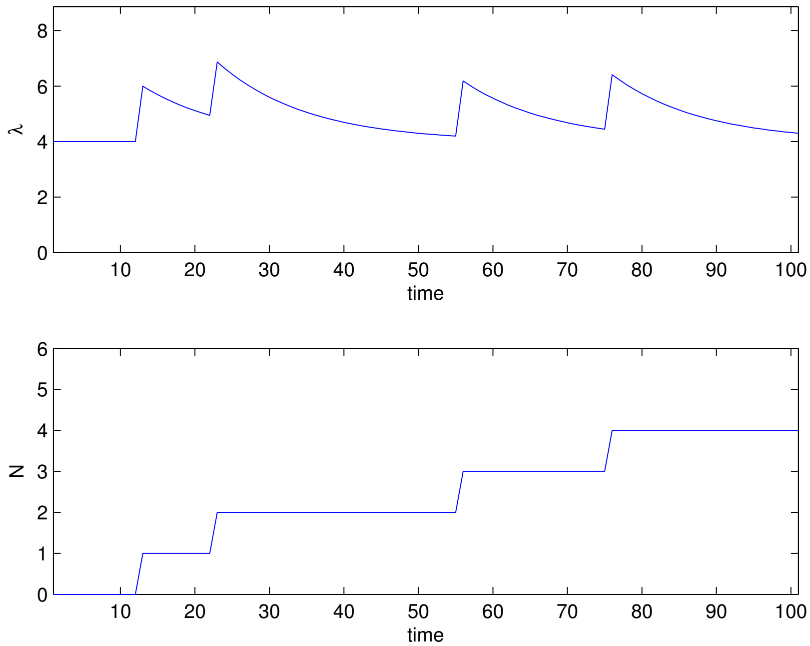


Figure 11: Sample trajectory from a time-stepped Hawkes process simulation.

of multiple events in a single time step. Shown in figure 11 is a sample trajectory from this scheme - it can be seen that λ has a smooth curve.

In the conventional method for pricing tranches, it is important to factor into consideration the exact times at which losses occur and the corresponding discounting. Since formula 1 refers to the cumulative loss at time T, we shift a loss at time t_i forward to time T by multiplying with $e^{r(T-t_i)}$.

Code Listing

```
%Program to obtain tranche prices on earthquake claims

%Input file consists of four columns:
%time latitude longitude magnitude
f     = fopen('../data/eastla.dat');
data  = fscanf(f, '%f %f %f %f', [4 inf])';
fclose(f);

%Store time start
t0 = data(1,1);
%Remove first point as its not counted
data = data(2:size(data,1), :);
times = data(:,1) - t0;
lats  = data(:,2);
longs = data(:,3);
mags  = data(:,4);

%Thin data a little
ind = find( mags > 4 );
```

```
mags  = mags(ind);
times = times(ind);


npts  = length(times);

%Fit model to data
params = hawkesfit( times, mags )
nu    = params(1);
alpha = params(2);
beta  = params(3);

%Simulate and price
%Interest rate
r = 0.05 / 365;
T = 5 * 365;
ntraj = 10000;

%Loss distribution
L = hawkesim(nu, alpha, beta, T, mags, r, ntraj);

%The nominal is decided based on the expected losses and some margin
%possibly chosen by VaR requirements by law
%The offset represents a 'deductible' for the insurance company
%The tranche prices obtained are arbitrary in the sense that one may
%choose different offset/nominal and structure the tranches accordingly
mL = mean(L);
offset  = mL - 3 * std(L);
nominal = (mL - offset) + 3 * std(L);


%Reference process
%From the reinsurer's perspective, it is some market portfolio
mu    = 0.15 / 365;
sigma = 0.2 / sqrt(365);
Hinvestor = exp( (mu - 0.5 * sigma^2) * T + sigma * sqrt(T) * randn(ntraj, 1) );


%From the company's perspective
eta  = 0.3 * mL;
phi = ( mL + 3 * std(L) ) / T;
Hcompany = phi / mu *( exp(mu * T) - 1 ) + eta * exp(mu * T) - L;

%Find prices for all tranches in one go

Ktranche = [0 3 7 10 15 30 100] / 100;
nK = length(Ktranche);
pcompany = zeros(nK - 1, 1);
pinvestor = zeros(nK - 1, 1);
for i = 1 : nK - 1
    pcompany(i) = exp(-mu * T) * mean( uprime(Hcompany) .* trancheloss( (L - offset) / nominal, Ktranche(i), Ktranche(i+1)) ) /
    pinvestor(i) = exp(-mu * T) * mean( trancheloss( (L - offset) / nominal, Ktranche(i), Ktranche(i+1) ) );
end
pcompany
pinvestor


%Returns the maximum likelihood estimators for a Hawkes process
%given time and magnitude data
function p = hawkesfit(t, mag, nu0, alpha0, beta0)
    %Initial guess for nu, alpha, beta:
    N      = length(mag);
    if nargin == 2
        nu0    = N / t(N);
        alpha0 = nu0;
```

```
        beta0  = 10;
    end

    %p = fmincon( @(x) -hpll( x(1), x(2), x(3), t, mag ), [nu0 alpha0 beta0], [], [], [], [], [0 0 0], [inf inf inf] );
    p = fminsearch( @(x) -hpll( x(1), x(2), x(3), t, mag), [nu0 alpha0 beta0] );
end

%Log likelihood function for Hawkes process
%Given parameters nu, alpha and beta, computes the Loglikelihood for the
%intensity process given by
%
% lambda(nu, alpha, beta)(t) = nu + \sum_{t_i < t} alpha e^{-\beta*(t-t_i)}

function out = hpll( nu, alpha, beta, t, mag )
    N = length(mag);

    %Explicit loop needed for numerical reasons
    %Maybe be fixable if beta is small.
    lambda = 0;
    sum_loglambda = log( nu + eps );
    for i = 2 : N
        %This is lambda without nu
        lambda = (lambda + alpha * mag(i-1) )* exp( -beta * (t(i) - t(i-1)) );
        sum_loglambda = sum_loglambda + log( lambda + nu + eps );
    end

    out = -nu * t(N) - alpha / beta * sum( mag .* (1 - exp(-beta * (t(N) - t)) ) ) + sum_loglambda;

end



%exponentially decaying feedback Hawkes process simulation
%as a cluster poisson process

%References:
%Moller and Rasmussen,
%Lewis and Schedler

%Using thinning algorithm for Hawkes process
%This is marginally slower than the adaptive timestep method
%but has no approximation tradeoffs

%Returns a loss distribution L_T
% nu, alpha, beta - Hawkes parameters
% T   - horizon
% mag - empirical distribution of magnitudes
% r   - interest rate
function L = hawkesim(nu, alpha, beta, T, mag, r, ntraj)
L = zeros(ntraj, 1);
t = zeros(ntraj, 1);
lambda    = nu * ones(ntraj, 1);
lambdabar = lambda;
arrivals  = [];
counts    = zeros(ntraj, 1);
while any( t < T )
    %Simulate next arrival times
    dt = exprnd( 1./ lambdabar, ntraj, 1);
    t  = t + dt;

    %Only consider those that are within T
    ind  = find( t < T );
    nind = length(ind);

    %To avoid some dimension mismatch errors
```

```
    if nind == 0
        continue
    end

    %Update lambda to this time
    lambda(ind) = nu + (lambda(ind) - nu) .* exp(-beta * dt(ind));

    %accept with prob lambda / lambdabar
    ind  = ind( rand(nind, 1) < lambda(ind) ./ lambdabar(ind) );
    nind = length(ind);

    %To avoid some dimension mismatch errors
    if nind == 0
        continue
    end



    %For the ones that we accept, increment lambda and lambdabar
    lambda(ind) = lambda(ind) + alpha;
    lambdabar(ind) = lambda(ind);
    %Generate magnitudes for these trajectories
    m = randsample(mag, nind, 1);
    %Add to cumulative loss with time shift
    L(ind) = L(ind) + lossfunc(m) .* exp( r * (T - t(ind)) );
end

end

%Derivative of utility function at x
%log utility:
function y = uprime(x)
    y = 1./x;
end

function y = trancheloss(x, Klower, Kupper)
    y = max( x - Klower, 0 ) - max( x - Kupper, 0);
end
```

References

[1] Davis M. H. A. *Mathematics of Derivative Securities.* Cambridge University Press, 1998. pp. 216-218.

[2] United States Geological Survey. Online data.

[3] Lewis P. A. W. and Shedler G. S. Simulation methods for poisson processes in nonstationary systems. In *Proceedings of the 10th conference on Winter simulation - Volume 1.* IEEE Press, Piscataway, NJ, USA.