ARCH 2016.2

May 16, 2016

Shannon Collier, Shiyao Li

SHORECREST HIGH SCHOOL

Project Math Minds

Analyzing Risk and Reward in California Earthquake Insurance

Shannon Collier & Shiyao Li

2/14/2016

Table of Contents

1. Executive Summary

This paper details the construction of a mathematical model to estimate the losses due to earthquakes in California. Each type of earthquake (moderate, strong, and major) was modeled as a separate binomial function and then multiplied by the appropriate amount of damages, considering the number of households they would affect and the severity of damage done. Using this model, we simulated 1,000 years of earthquakes to develop a distribution of insured earthquake losses in California, and then used these trials to draw conclusions. Over 89% of the trials showed a positive net revenue when the premiums were factored in, but the majority of the rest of the years had losses ranging around $1.5 billion. For the NECF to be 95% confident they could pay all losses, they would require reserves of approximately $1.33 billion; being 99% confident would require reserves of approximately $1.53 billion.

However, it should be noted that the current model is on the simplistic side in several regards and thus may not be fully accurate. We proposed five possible enhancements to the model, each of which would require additional research but would likely add accuracy. They consisted of considering how large earthquakes affect the occurrence of smaller ones, eliminating the threefold classification of earthquakes in favor of only using magnitude, creating a more accurate damage multiplier, eliminating the four-fold classification of premiums and claims in favor of a more nuanced understanding of both, and adding in region as a factor.

Using the Large Earthquake data, we developed an exponential regression equation relating magnitude of the earthquake to the overall damages it caused in billions of dollars; and we then modified the Basic Damage Multiplier to render the two comparable. The new equation for Large Earthquakes has a much steeper slope than the Basic Damage Multipliers, suggesting that the latter may not be a stellar way of predicting actual damages due to earthquakes.

We also had several suggestions for how to increase the percentage of households with earthquake insurance. The simplest one is to simply being a campaign educating homeowners about the severity of the risk earthquakes pose and the importance of being insured for them, but we also considered lobbying with lawmakers to either begin requiring earthquake insurance (to align it with automobile and some other kinds of insurance) or to create financial incentives for insurance companies or homeowners.

2. Introduction

Although few people ever worry about their occurrence, earthquakes pose a considerable risk to people living on the west coast of the United States. Earthquakes cause enormous amounts of damage whenever they occur, both in terms of lives but especially in terms of property damage; yet they vary widely in terms of frequency and severity, making them difficult for homeowners or insurance companies to predict. To combat this uncertainty, the National Earthquake Catastrophe Fund - a hypothetical organization offering earthquake insurance - desires to know the distribution of insured losses in California in a year. This is the task Project Math Minds set out, and that this paper addresses: to predict the insured damages caused by earthquakes in a year. The project introduces the idea of using mathematical models to calculate the probability of unlikely events and the importance of thinking precisely. Based on research and the assumptions provided, we developed a complicated probability model to estimate the potential losses that the NECF might bear in California next year. Each of the three types of earthquakes (moderate, strong, and manor) are separately modeled as binomial functions, and we then take into account the number of households each type hits and how much damage is typically done to per household. All of this is combined into one model, which we used to simulate 1,000 years. Finally, we drew conclusions based on those trials.

3. Body

3.1. Methods

We created mathematical models for the number, type, and cost of earthquakes that could occur in any random year. There are three types of earthquakes we are concerned with - moderate, strong, and major - each of which encompasses a range of magnitudes. Any earthquake with a magnitude between 5.0 and 5.9 is classified as moderate, while strong encompasses 6.0-6.9 and major encompasses 7.0-7.0. In reality, there is a substantial difference between a magnitude 6.0 earthquake and a magnitude 6.9, but to simplify the model we assumed that all earthquakes of each type are the middle of their range (moderate, 5.5; strong, 6.5; and major, 7.5).

Each type of earthquake can be modeled as a different binomial random variable expressing the number of times they occur per year. By multiplying those random variables by expected damages per earthquake and then combining all three types of earthquake, we created a three-part probability model for the damages in a random year. Three randomly generated numbers can thus together be used to simulate one year.

3.1.1. Binomial Distributions

3.1.1.1. Why Binomial Functions?

For each earthquake, the number of earthquakes that occur per year is a binomial random variable, because they fulfill the four requirements for a binomial setting.
Requirements for a Binomial Setting

- Each observation falls into one of two categories ('success' or 'failure'). Each month, either an earthquake occurs or it does not.
- There is a fixed number *n* of observations. Each year has exactly 12 months, and we are told to assume that each type of earthquake can occur only once per month.
- Each of *n* observations is independent. We are told to assume that all earthquake events are independent.
- The probability of success *p* is the same for each observation. All earthquake events are independent, and we assume that time of the year does not affect the probability of an earthquake occurring.

3.1.1.2. Binomial Models for each Earthquake

Binomial functions are determined by two numbers, the *n* number of trials and the *p* probability of success.

For each earthquake, *n* = 12, but the probability *p* that each kind of earthquake occurs is different.

Moderate earthquakes occur 3-4 times per year, which we rounded to 3.5 times per year.
(3.5 moderate earthquakes per year) x (1 year per 12 months) = 0.291666 moderate earthquakes per month
n = 12
p = 0.291666

Strong earthquakes occur once every 2-3 years, which we rounded to once every 2.5 years.
(1 strong earthquake per 2.5 years) x (1 year per 12 months) = 0.0333 strong earthquakes per month
n = 12
p = 0.0333

Major earthquakes occur once every 10 years.
(1 major earthquake per 10 years) x (1 year per 12 months) = 0.008333 major earthquakes per month
n = 12
p = 0.008333

3.1.2. Damages per Earthquake

According to the US Census Bureau, there were 12,542,460 households in California from 2009-2013, so this is the number we have based our calculations on (“California”). The rest of the statistics were provided by The Actuarial Foundation and their instructions for Project Math Minds.

The damage an earthquake will cost depends on the number of households affected and the average cost per insured household for that type of earthquake. In turn, the cost per insured household is affected by the type of household and the magnitude of the earthquake in question. Rather than separately calculate the damages per type of household (ie the damage to homes, to renters, to condos, etc.), we calculated the weighted average claim payment for an average insured household. This was done by multiplying the average claim payment per type of household by that type of household’s proportion of all insured households.

[($11,500 in damage to a homeowner) x (42% of insured households are homeowners)] + [($2,000 in damage to a mobile home owner) x (16% of insured households are mobile home owners)] + [($8,500 in damage to a condo unit owner) x (15% of insured households are condo unit owners)] + [($1,500 in damage to a renter) x (27% of insured households are renters)] =
$6,830 in damage to an average insured household in a moderate earthquake

That is the average damage to an insured household in a moderate earthquake. However, the magnitude of the earthquake also affects its damages, because higher magnitude earthquakes cause more structural damage than lower magnitude earthquakes. The damage multiplier function is as follows:

Damage = average claim payment from a moderate earthquake x
(10^[magnitude – 5.5])/3

This is all of the information needed to calculate the overall damages caused by each type of earthquake.

3.1.2.1. Damages per Moderate Earthquake

Only 7% of California households have earthquake insurance, and of those 7%, 0.5% will be affected by any given moderate earthquake.

(12,542,460 households) x (0.07 with earthquake insurance) x (0.005 hit by a moderate earthquake) = 4390 insured households hit by a moderate earthquake

We do not need to apply the damage multiplier for moderate earthquakes because the given average claim payments are for a magnitude 5.5 earthquake.

(4390 insured households hit by a moderate earthquake) x ($6,830 in damage to an average insured household in a moderate earthquake) = $29,983,700 in overall damage caused by a moderate earthquake

3.1.2.2. Damages per Strong Earthquake

A strong earthquake affects 0.7% of California households.

(12,542,460 households) x (0.07 with earthquake insurance) x (0.007 hit by a moderate earthquake) = 6102 insured households hit by a strong earthquake

($6,830 in damage to an average insured household in a moderate earthquake) x (10 ^ [6.5-5.5]) / 3 = $22,766.67 in damage to an average insured household in a strong earthquake

(6102 insured households hit by a strong earthquake) x ($22,766.67 in damage to an average insured household in a strong earthquake) = $138,922,200 in overall damage caused by a strong earthquake

3.1.2.3. Damages per Major Earthquake

A major earthquake affects 0.8% of California households.

(12,542,460 households) x (0.07 with earthquake insurance) x (0.008 hit by a major earthquake) = 6973 insured households hit by a major earthquake

($6,830 in damage to an average insured household in a moderate earthquake) x (10 ^ [7.5-5.5]) / 3 = $227,666.67 in damage to an average insured household in a major earthquake

(6973 insured households hit by a strong earthquake) x ($227,666.67 in damage to an average insured household in a strong earthquake) = $1,587,519,690 in overall damage caused by a major earthquake

3.2. Results

3.2.1. Simulated Trials

To view the simulated trials, look in the “Year Trials” worksheet in the attached Excel spreadsheet.

3.2.2. Basic Problem Answers

3.2.2.1. Develop a distribution of expected losses based on simulated results. How many trials did you include and why?

Ultimately, we included 1,000 trials. We began with only 200 trials, but with only 200 trials each time a new set of random numbers was generated the distribution of net revenue changed dramatically. This meant that 200 trials would not be enough to establish a trustworthy distribution. To solve this problem, we incrementally increased the number of trials used until the distribution seemed to no longer vary significantly between sets of random numbers. That number of trials was 1,000.

3.2.2.2. How much reserves are required to be 95% certain you can pay all losses?

Each year, the required reserves can be found by subtracting the amount of losses from the revenue gained from annual premiums. But, to be 95% certain the insurance company could pay all losses, they would need to have enough reserves to pay all losses in 95% of years (because by wanting only 95% certainty, they have decided to ignore the worst 5% of years). In order to be sure they can pay for that 95% of the time, they would need enough reserves to pay for the worst-case scenario of those years. So, to figure out what that required reserve would be, we sorted our trials from by reserves from least to greatest. Then, we found trial 950 - the worst possible year that occurred in the best 95% of the time - and that year needed a reserve of $1,329,410,872.37. Thus, in our model, the insurance company would need a reserve of $1,329,410,872.37 to be 95% certain they could pay all losses.

3.2.2.3. How much reserves are required to be 99% certain you can pay all losses?

We used the same process for 99% certainty that we used for 95% certainty. The company needs to be able to pay for the worst possible year that occurs in the best 99% of the time, which in our simulation was trial 990. The required reserves for trial 990 was $1,529,295,876.57, so that is the reserves the company should have to be 99% certain they can pay all losses.

3.2.2.4. Do the premiums charged appear reasonable relative to the expected losses? Why or why not?

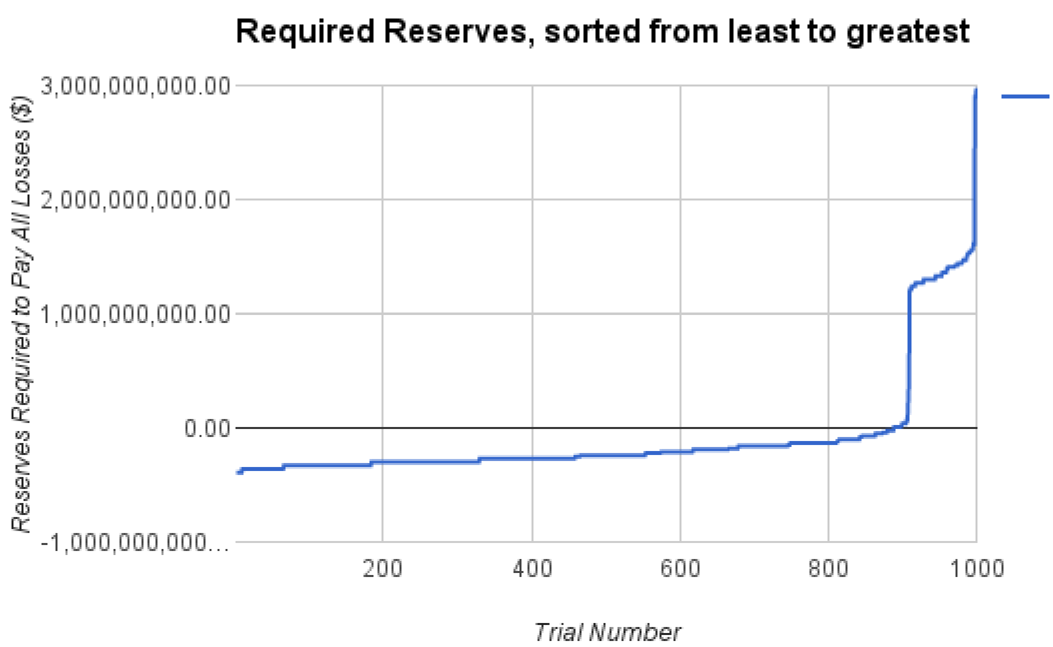
The premiums appear expensive relative to the expected losses. The expected losses per year is -$321,000,000, while the expected gains from premiums is $390,000,00. That means the insurance company will, in an average year, have a net revenue of $69,000,000, which seems very high.

3.2.2.5. What other factors may influence the premium amounts?

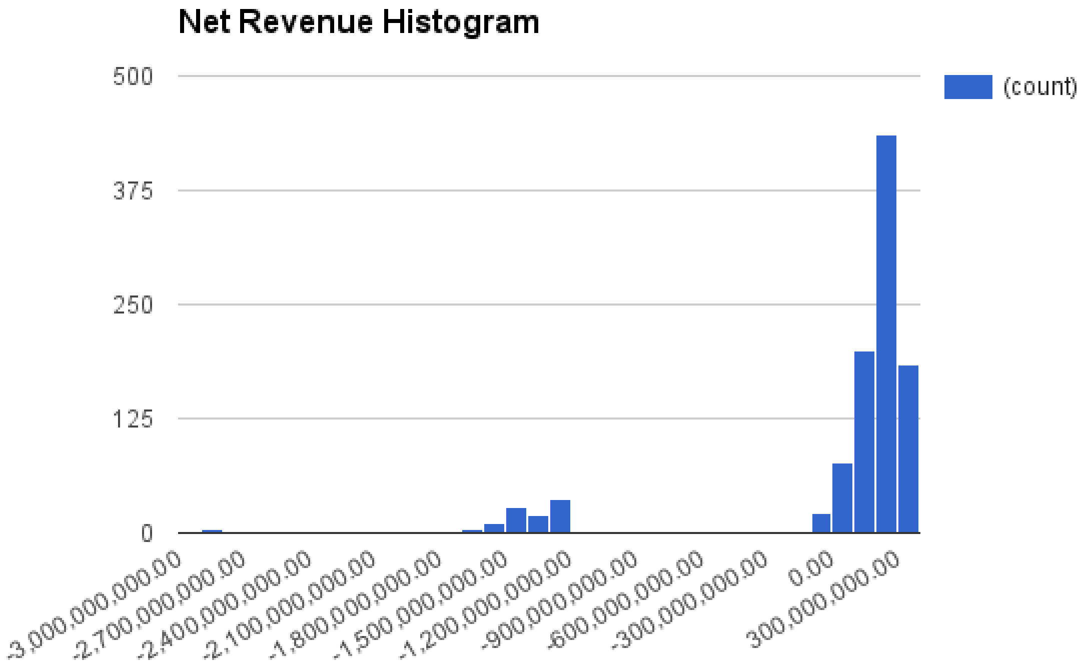
In reality, premiums vary widely among different households. Households in more earthquake-prone areas face higher premiums, as do households that are older or not built to be very earthquake-resistant, because all the of those factors increase the risk to the insurance company ("Know"). Additionally, as with most types of insurance, different earthquake insurance plans have different premiums: higher deductible, lower premium, and vice versa.

3.2.2.6. Please provide a visual display of your work.

The following graph is from the simulated trials. The reserves we defined as the difference between the loss and revenue. In other word, the reserves are the opposite of net revenue.



The following chart is the histogram of the net revenue (revenue minus loss) from our 1000 simulated trials.



3.2.3. Advanced Problem Answers

3.2.3.1. What five enhancements would you consider making to the model to make it more accurate?

Geologically speaking, earthquake events in one region within a month are not completely independent, as larger earthquakes are often foreshadowed or followed by smaller earthquakes. According to a study by Stanford, “conventional seismic earthquakes follow a statistical pattern: For every step down in magnitude, there is a 10-fold increase in the number of smaller earthquakes. For example, for every magnitude 6 there will be 10 magnitude 5s, 100 magnitude 4s and so on. Likewise, if there is one M6 per year, it will correlate with one M7 per decade and one M8 every 100 years.” (”Silent”) The model could thus be enhanced by analyzing the past frequencies of smaller earthquakes before and after historically large earthquakes, and determining whether there is a statistical relationship between the two. If so, the model could be enhanced by taking the relationship between large earthquakes and small earthquakes into account on a month-by-month basis. Mathematically, this relationship would likely be expressed through conditional probabilities, where the likelihood of a smaller earthquake each month is higher or lower depending on whether a large earthquake has occurred.

The current model assumes that earthquakes come in three different magnitudes - 5.5, 6.5, and 7.5 - which is an oversimplification. The model would be more accurate if it was solely based on magnitude, rather than a ‘type’ of earthquake, because as noted previously, there is a large difference between a magnitude 5.0 and a 5.9 earthquake. Creating this enhancement would involve directly, mathematically relating the magnitude of an earthquake to the number of households it affects, rather than simply classing them into three categories.

The current model does contain a damage multiplier based on magnitude, but this multiplier does not seem to relate magnitude of earthquake to overall insured loss very well, as it does not match the historical data provided (see section 3.2.3.5). The model could thus be enhanced by creating a more accurate damage multiplier. It should be noted that the damage multiplier we created in section 3.2.3.5 would not be appropriate to use for this purpose, because that predicts overall damages, not the damages to one household.

The second shortcoming of the current model’s simplistic classification of premiums and losses only contains four categories (homeowners, mobile home owner, condo and renter). In reality, the claim payment is related to the premium, and is not solely determined by the type of household. The premium amount is determined by the value of the house. The model can be better enhanced if we use a more accurate premium / claim payment calculation.

The current model tries to predict dollar losses based solely on the Magnitude Scale measurement of earthquakes, but ignores where the earthquake occurs. Earthquakes with higher magnitudes are larger and can cause more damage but more than just the Magnitude Scale may affect the the loss in terms of dollar amounts. The damages caused by an earthquake are also related to its depth and its location. For example, in 1999, a magnitude 7.1 earthquake struck in CA and caused negligible damage, while a magnitude 6.9 earthquake in 1989 caused $6 billion in property damage (“Major”). This massive disparity is due to the location of these earthquakes: one struck in uninhabited hills, while the other hit Loma Prieta. To better enhance the accuracy of the model, we can examine the damages in different regions and <u>add region in as a consideration</u> in the model.

3.2.3.2. What is the expected amount of uninsured losses?

3.2.3.2.1. Expected Uninsured Losses due to Moderate Earthquakes

Expected number of uninsured households hit by one moderate earthquake: (12,542,460 households in CA) x (1 - (0.07 with insurance)) x (0.005 hit by the moderate earthquake) = 58,322 uninsured households hit by a moderate earthquake

Expected damage to an average household caused by a moderate earthquake: $6830

Expected uninsured losses from one moderate earthquake: (58,322 households hit) x ($6830 per household damaged) = $398,339,260

Expected uninsured losses due to moderate earthquakes in an average year: (3.5 moderate earthquakes per year) x ($398,339,260 lost per moderate earthquake) = $1,394,187,410

3.2.3.2.2. Expected Uninsured Losses due to Strong Earthquakes

Expected number of uninsured households hit by one strong earthquake: (12,542,460 households in CA) x (1 - (0.07 with insurance)) x (0.007 hit by the strong earthquake) = 81,651 uninsured households hit by a strong earthquake

Expected damage to an average household caused by a strong earthquake: ($6830, the weighted average claim payment) x (10 x ^ (6.5-5.5) / 3) = $22,766.67

Expected uninsured losses for all households from one strong earthquake: (81,651 households hit) x ($22,766.67 per household damaged) = $1,858,921,372

Expected uninsured losses for all households due to strong earthquakes in an average year: (1 strong earthquake per 2.5 years) x ($1,858,921,372 lost per strong earthquake) = $743,568,549

3.2.3.2.3. Expected Uninsured Losses due to Major Earthquakes

Expected number of uninsured households hit by one major earthquake: (12,542,460 households in CA) x (1 - (0.07 with insurance)) x (0.008 hit by the major earthquake) = 93,316 uninsured households hit by a major earthquake

Expected damage to an average household caused by a major earthquake: ($6830, the weighted average claim payment) x (10 x ^ (7.5-5.5) / 3) = $227,666.67

Expected uninsured losses for all households from one major earthquake: (93,316 households hit) x ($227,666.67 per household damaged) = $2,124,492,076

Expected uninsured losses for all households due to major earthquakes in an average year: (1 moderate earthquakes per 10 years) x ($2,124,492,076 lost per moderate earthquake) = $212,449,207.6

3.2.3.2.4. Expected Total Uninsured Losses

The total expected uninsured losses in a year is the combination of uninsured losses from each type of earthquake.

($1,394,187,410 expected uninsured losses to moderate earthquakes) + ($743,568,549 expected uninsured losses to strong earthquakes) + ($212,449,207.6 expected uninsured losses to major earthquakes) = $2,350,205,167

3.2.3.3. List any assumptions you would make and provide support for them.

In order to determine the expected amount of uninsured losses using the information we were given, we need to know the distribution of household types among uninsured households. We were given the distribution of household types among insured households, but it is possible that some types of households are overrepresented respective to the total California population. However, we do not know the distribution of insured households, so to answer this question we assumed that insured households and uninsured households have the same distribution of household types.

3.2.3.4. What steps would you consider to increase the percentage of households with earthquake insurance?

Historically, it has proven challenging to increase the rates of earthquake insurance among households. However, there are several possible steps that could be taken to increase these percentages, which break into one of two main categories: convincing homeowners to purchase earthquake insurance, and convincing insurance companies to lower their rates.

Many homeowners do not own earthquake insurance because they do not understand how much of a danger earthquakes present. This could potentially be remedied by advertising the frequency of earthquakes and the severity of the risk they pose.

Business and government policies are related in many means. In order to increase insurance rates, insurance companies can lobby with lawmakers to make earthquake insurance a requirement, just as automobile insurance is required. Another route with lawmakers would be providing financial incentives, either for insurance companies to offer it or for homeowners to purchase it.

Considering nature of supply and demand, the companies can lower the premium price in order to increase the insurance rate. Especially in the rural areas, where the damages of an earthquake is not as huge as in downtown areas, the homeowners are less likely to purchase insurance. Lowering the premium will attract many homeowners. Direct subsidies to the insurance companies would cut their costs and allow them to lower premiums, as would tax deductions. Or, the government could incentivize homeowners to purchase earthquake insurance.

3.2.3.5. Based on the data provided below for Large Earthquake Events, develop a mathematical function for the damage multiplier. How does this compare with the damage multiplier used for the Basic Problem?

We built an exponential regression based on the data for the Large Earthquake Events.

Large Earthquake Damage Multiplier: 1.6285764728872x10^-5 x 7.4869221022039^(magnitude) = overall damages in billions of dollars

However, comparing this function to the Basic Damage Multiplier is difficult, because this new regression calculates the expected overall damages in billions of dollars, whereas the Basic Damage Multiplier calculates the expected damages to one household in dollars. So, to compare, we've modified the Basic Damage Multiplier to be comparable.

Basic Damage Multiplier: ($6830, the weighted average claim payment) x (10 x ^ (magnitude -5.5) / 3) = expected damages per household in dollars

Modified Basic Damage Multiplier: (damages per household) * (number of households affected) / (1,000,000,000)

For 5.0 < magnitude < 5.9

(Damages per household) * (CA households * 0.005) / (1,000,000,00)

[($6830, the weighted average claim payment) x [12,542,460 households * 0.005 hit] x (1/1,000,000,000) = expected overall damages in billions of dollars

For 6.0 < magnitude < 6.9

(Damages per household) * (CA households * 0.007) / (1,000,000,00)

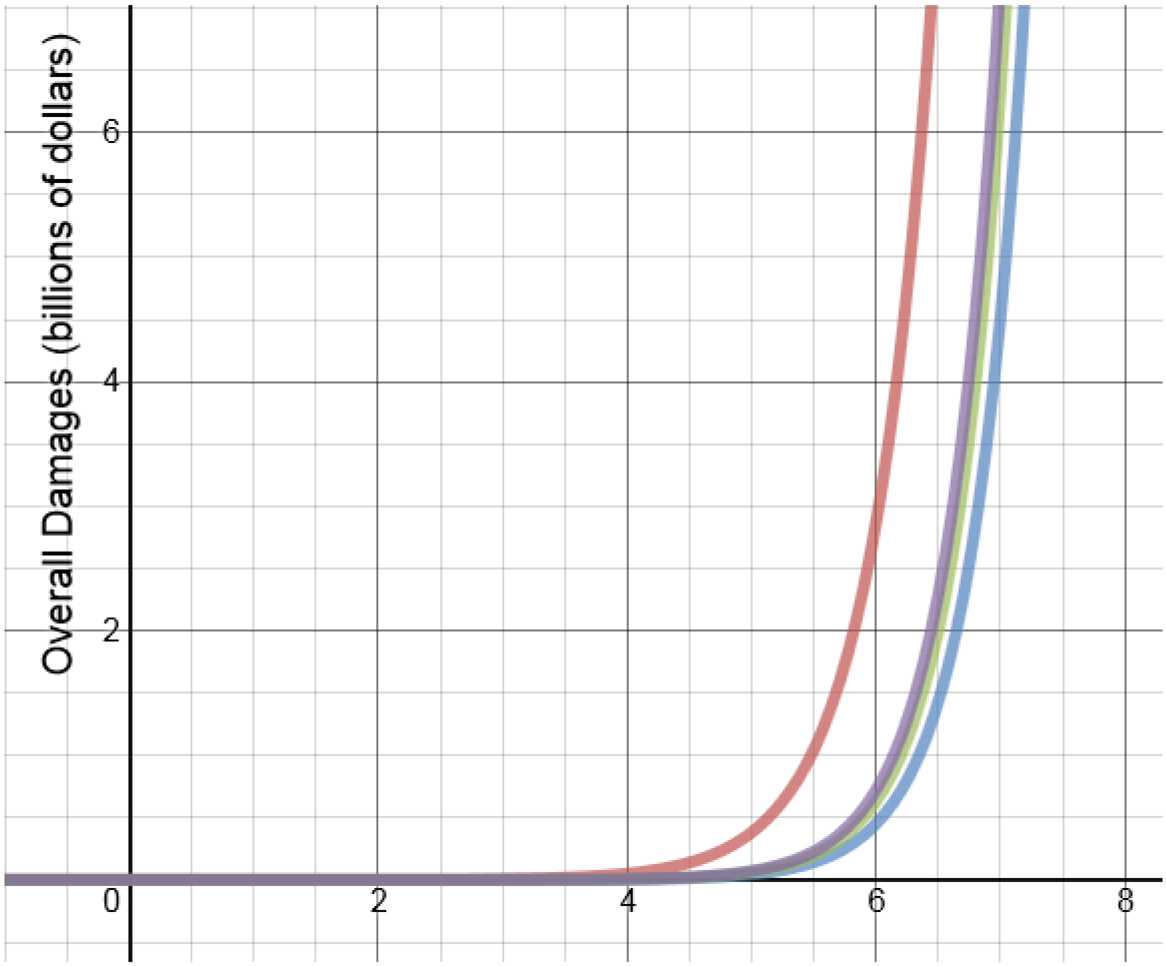
[($6830, the weighted average claim payment) x(10 ^ (magnitude -5.5) / 3) x [12,542,460 households * 0.007 hit] x (1/1,000,000,000) = expected overall damages in billions of dollars

For 7.0 < magnitude < 7.9

(Damages per household) * (CA households * 0.008) / (1,000,000,00)

[($6830, the weighted average claim payment) x(10 x ^ (magnitude -5.5) / 3) x [12,542,460 households * 0.008 hit] x (1/1,000,000,000) = expected overall damages in billions of dollars

Below is a visual comparison of the Large Earthquake Damage Multiplier with the three different Modified Basic Damage Multipliers, relating the overall damages in billions of dollars to the magnitude of an earthquake. The red line is the Large Earthquake Damage Multiplier, while the blue, green, and purple lines represent the Modified Basic Damage Multipliers for moderate, strong, and major earthquakes, respectively. As is clearly visible, the Large Earthquake Damage Multiplier has a much steeper curve than the Modified Basic Damage Multipliers.



4. Conclusions

Based on the 1,000 trials from our mathematical model, for the NECF to be 95% confident they could pay all losses, they would require reserves of approximately $1.33 billion; being 99% confident would require reserves of approximately $1.53 billion. However, it should be noted that the current model and its damage multiplier is on the simplistic side in several regards and thus may not be fully accurate. Indeed, when compared to the exponential regression equation developed to fit the historical data on Large Earthquake Events, it was clear that the Basic Damage Multiplier does not accurately map magnitude of earthquake to the overall damages caused. Thus, we proposed five possible enhancements to the model, each of which would require additional research but would likely add accuracy. They consisted of considering how large earthquakes affect the occurrence of smaller ones, eliminating the threefold classification of earthquakes in favor of only using magnitude, creating a more accurate damage multiplier, eliminating the four-fold classification of premiums and claims in favor of a more nuanced understanding of both, and adding in region as a factor.

We also had several suggestions for how to increase the percentage of households with earthquake insurance. The simplest one is to simply being a campaign educating homeowners about the severity of the risk earthquakes pose and the importance of being insured for them, but we also considered lobbying with lawmakers to either begin requiring earthquake insurance (to align it with automobile and some other kinds of insurance) or to create financial incentives for insurance companies or homeowners.

References

"California QuickFacts from the US Census Bureau." *California QuickFacts from the US Census Bureau*. N.p., n.d. Web. 14 Feb. 2016.

“Know Your Risk for Earthquake Damage.” *California Earthquake Authority*, 9 Sept 14. Web. 11 Feb 2016.

“Major California Earthquakes by Month”. *Century-National*. Century National, n.d. Web. 13 Feb 2016.

“Project Math Minds Assignment.” *The Actuarial Foundation*, Liberty Mutual Insurance. N.p., n.d. Web. Feb 10, 2016.

"‘Silent Earthquakes’ May Foreshadow More Destructive Temblors, Study Finds." *Stanford University*. N.p., n.d. Web. 13 Feb. 2016.