Modeling of Catastrophes and Estimation of National Catastrophe Fund

Honors Research Thesis

Presented in partial fulfillment of the requirements for graduation with honors research distinction in Actuarial Science in the undergraduate colleges of The Ohio State University

by

Nitish Aggarwal

The Ohio State University

April 2013

Project Advisor: Professor Chunsheng Ban, Department of Mathematics - Actuarial Science

ABSTRACT

Bill 'H.R. 2555: Homeowners' Defense Act of 2010' was considered by the United States of America (US) Congress to form National Catastrophe Risk Consortium. The bill was introduced on April 27, 2010, but was not enacted. One of the functions of the consortium was to fund a National Catastrophe Fund to help public and insurance companies meet the liability claims from hurricane, fire, and blizzard. The bill has to pass through House Financial Services Committee which takes into account the projected costs, disbursements, and the amount required to be appropriated for the task and its source. In this project the following are accomplished:

Part I: Modeling

- Employed the method of Maximum Likelihood Estimation (MLE) to estimate the parameters such as mean of count of fires, hurricanes and blizzards.
- Modeled the number of acres burnt by fires by Weibull distribution.
- Modeled the economic damages caused by blizzards by Poisson Distribution.
- Employed Poisson distribution to model the number of hurricanes occurring in each hurricane category.

Part II: Estimating

- Estimated 2013 claims for hurricane, fire, and blizzard in United States of America using data for three catastrophes for the last 100 years.
- Carried out Linear regression analysis to compare the results of the new model for acres burnt from fires and economic damage from blizzards.
- Predicted the economic damages from the three disasters in 2013 to be 46.7 billion dollars.
- The National Catastrophe Fund for the liabilities of fire, blizzard, and hurricane needs to be funded by 10% of the estimated economic damages which for 2013 amounts to 4.67 billion dollars.

Background

Natural disasters such as earthquakes, tsunamis, wildfires, hurricanes, and tornadoes have the capability to cause human mortality as well as enormous economical damage. The National Fire Protection Association gives an overview of economic damage from fires in the US[1]. Brent Weisman, a University of South Florida professor,[2] provides basic statistics for major hurricanes in Florida. Fahy et al.[3] gives the statistics for fire damage in the US in year 2010. Hamid et al.[4] have characterized damage from public hurricane loss model for Florida and have carried out various simulations. Recovery and mitigation efforts use a lot of time and economic resources to ensure safety of citizens[5]. Richard Anthnes[6] utilizes the generation of inertia-gravity waves caused by the imbalances in the data. This technique accounts for stronger hurricanes that skew the data. Henry Wright[7] describes effects of fire on the land that is used for agricultural purposes, specifically in California. Scholars have described most of the factors that affect damage, however, in the literature there is an absence of economic loss predictions based upon historic data.

Significance of the project

Insurance companies by and large compensate the public for property damage after the disasters. Sometimes their funds are insufficient to cover the damages. US congress had proposed a National Catastrophe Insurance Fund to help the insurance companies or the public to tide over these damages. This would help in keeping a check on rising insurance rates. It considered a bill `H.R. 2555: Homeowners' Defense Act of 2010'[13] to form National Catastrophe Insurance Fund (NCIF). Before such a policy is enacted into law by US Congress, the bill has to pass through the House Financial Services Committee which takes into account the projected costs, disbursements and the amount required to be appropriated for the task and its source. In this project the following are attempted for the first time:

- **Application of Poisson Distribution in predicting economic damages for blizzards.**
- **Modeled the number of acres burnt using Weibull Distribution.**
- **Predicting the parameters for the count of fires and blizzards by Maximum Likelihood Estimation method.**

Objective

The primary objective was to estimate the economic damages from hurricane, fire, and blizzard in United States of America for 2013 using data for the three catastrophes from the last 100 years and the amount required to fund National Catastrophe Insurance Fund based upon the 10% liability. The estimates will be based upon mathematical models and application of statistics to data on these disasters for last hundred years. In the process of estimating the damages, secondary objective of constructing mathematical models was created. The parameters for estimating count of fires and blizzards occurring in a year were estimated by Maximum Likelihood Estimation. The economic damages for 2013 are estimated based upon the catastrophe events following Poisson Distribution and Weibull Distribution. The results are then compared to those obtained from linear regression. For the 2013 year, based on annual data from 1900, an estimate of NCIF reserve fund is computed.

Materials (Data Source and Software)

Annual data of number of hurricanes, acres of areas burnt in fires and economic damage from blizzards from 1900 through 2008 was obtained from Society of Actuaries website. For the fires, the numeric data consisted of count of fires and acres burned. For the blizzards, the numeric data consisted of storm count and economic damages. For the hurricanes, the numeric data consisted of storm category, including tropical storms and sub-tropical storms, and economic damages. The sample data for the three catastrophes is given in Appendix Tables 1-3.

Microsoft Excel, SAS[14] and Stata[15] statistical package in conjunction with Java were used to estimate the damages.

Methodology and Results

Part I: Modeling

Maximum Likelihood Estimation

Models require estimation of parameters that best fit the data. MLE (Maximum Likelihood Estimation)[16] is a standard approach in statistics to test hypothesis and estimate parameters. MLE is the value that makes the observed data the most probable. The following describes the theory behind Maximum Likelihood Estimation, an approach used to estimate parameters in the paper.

Let $X_1, X_2, X_3, \ldots X_n$ have a joint density as following:

$$f_\theta(x_1, x_2, x_3, \ldots x_n) = f(x_1, x_2, x_3, \ldots x_n|\theta).$$

The likelihood of θ becomes the function: $lik(\theta) = f(x_1, x_2, x_3, \ldots x_n|\theta)$ if observed values $X_1 = x_1$, $X_2 = x_2$, $X_3 = x_3$, ... $X_n = x_n$, are given. In the end, MLE of θ is the value of θ that optimizes $lik(\theta)$. $lik(\theta)$ is the probability of observing the given data as a function of θ.

If X_i are independent and identically distributed random variables, then

$$lik(\theta) = \Pi^n_{i=1} f(x_1|\theta)$$

Applying the logarithm to both sides:

$$l(\theta) = \sum^n_{i=1} log(f(x_1|\theta))$$

The objective becomes to maximize $l(\theta)$ instead of $lik(\theta)$ because maximizing the latter is arduous.

The value of θ can be found by $\frac{dl}{d\theta} = 0$ and $\frac{d^2l}{d\theta^2} < 0$. The first part is the derivative which maximizes the likelihood. The latter part is the second derivative to test if the function is concave, because derivative of 0 for a concave function results in maximum and not minimum.

By continuing this process, the maximum likelihood estimate equals the conventional average of the data.[21] Therefore, by average provides the maximum likelihood estimate for the Poisson distributed data.

Proposition: It was determined by MLE that the average number of fires in a year is 104,647. The average number of blizzard storms in a year is 2,157.

Weibull Distribution

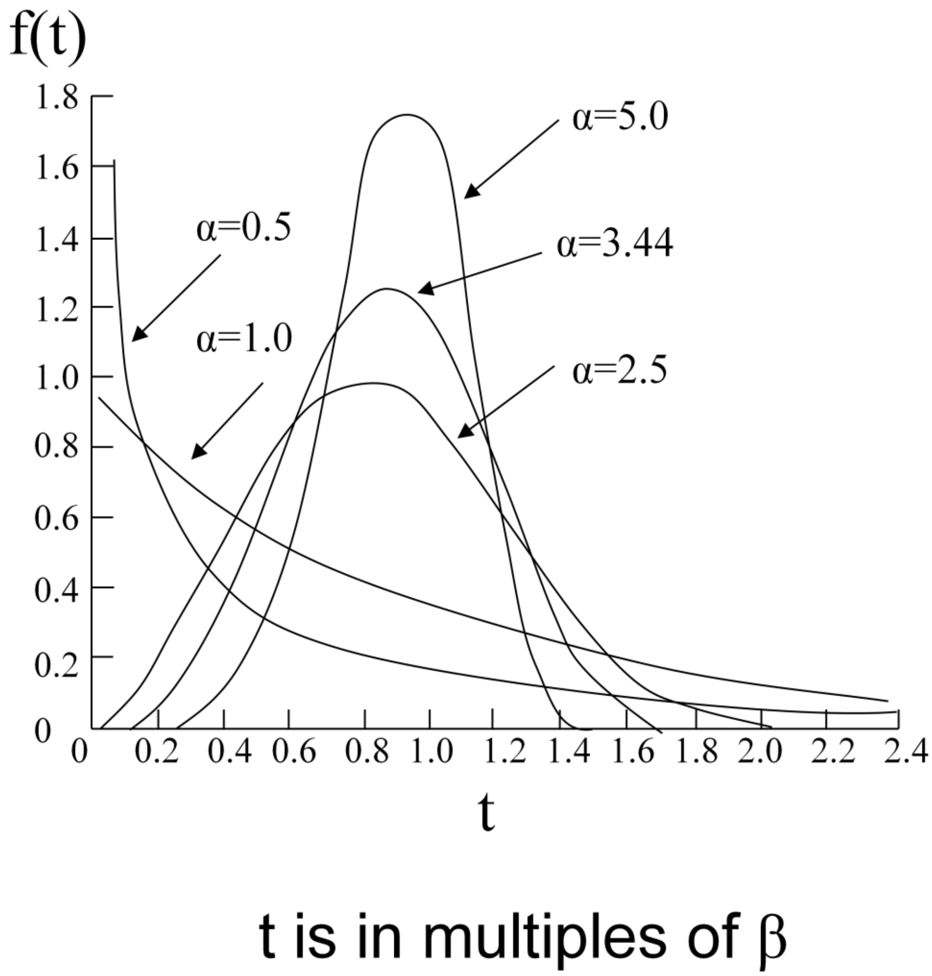
Weibull distribution was used to model the number of acres burned per year. Weibull distribution is a continuous probability distribution named after Waloddi Weibull.[11] It is a probability distribution with parameters α and β, where α > 0, and β > 0 is a Weibull distribution if the probability density unction of random variable x is:

$$P(x) = \alpha\beta^{-\alpha} x^{\alpha-1} e^{-(x/\beta)^{\alpha}} \qquad \text{for } x \geq 0$$

And 0 elsewhere.

α is the shape parameter and β is the scale parameter.

The shape of the distribution resembles that of the normal distribution as α increases. An example of this property is shown in Graph 1.



Graph 1. Probability Density function for Weibull Distribution[19]

A special property of Weibull Distribution is if $\alpha = 1$, then the Weibull distribution reduces to Exponential Distribution. Weibull Distribution is a recognized model for failure rate in time. If $\alpha = 1$, and hence the exponential distribution, then there is constant failure rate. If $\alpha > 1$, then there is an increasing failure rate, and decreasing failure rate is observed if $\alpha < 1$. The mean of the Weibull distribution can be found by integration x * p(x) over the domain and comes out to the following:

$$\mu = \beta\Gamma(1 + 1/\alpha)$$

where Γ is the gamma function.

Gamma function is a special function based on the factorial function over the non-negative integers. The formal definition of the gamma function is the following, for any positive integer z,

$$\Gamma(z) = \int_0^{\infty} x^{z-1} e^{-x}\, dx.$$

Essentially, the gamma function becomes $\Gamma(z) = (z-1)!$ for any positive integer z.

Proposition: The number of acres burned follows a Weibull Distribution with parameters α being 2.923 and β being 5.03×10^6. The mean for the distribution was calculated as 4.4866 million acres burned, and the standard deviation was 1.67×10^6 acres burned. With the inverse cumulative distribution function, it was determined that the 95th percentile of the acres burned were less than 7.32 million acres. The fitted Weibull distribution is shown in Appendix Figure 2. Appendix Figure 3 shows the P-P plot[17] or the probability plot that evaluates how close the data set concurs with the estimates from the modeled distribution. In the graph, the points do not stray from the y = x line, therefore the Weibull distribution is a good model.

Goodness of Fit tests

A goodness of fit test is used to test if a data sample came from a population with specific distribution. There are three main tests that are conducted: Chi-Square, Anderson-Darling, and Kolmogorov-Smirnov[12].

Chi-Square Goodness of Fit Test

Chi-Square goodness of fit test can be applied to any univariate distribution that has a cumulative distribution function. The data has to be in either class form (bins) or a histogram must be constructed to calculate the chi-square goodness of fit test. Another limitation of the test is that it requires adequate data set for the chi-square approximations to be accurate. One advantage for chi-square over the other two tests is that it can be applied to discrete distributions. The chi-square test is defined as followed:

Null Hypothesis: Data follows the specified distribution.

Alternate Hypothesis: Data does not follow the specified distribution.

Test Statistic: $$\chi^2 = \sum_{i=1}^{k}(O_i - E_i)^2/E_i$$

where k is the number of bins data is divided into. O_i and E_i are the observed frequency in ith bin and the expected frequency in the ith bin, respectively. There are several ways to calculate *k*, the number of bins, such as intervals with equal width. One empirical formulaic method to calculate *k* is the following: $k = 1 + \log_2 N$, where N is the sample size.

E_i can be calculated based on the cumulative distribution function, F(x), as following:

$$E_i = N(F(Y_u) - F(Y_l))$$

where N is the sample size, Y_u is the upper limit, and Y_l is the lower limit for ith class.

The chi-square test statistic is used to determine if the null hypothesis is rejected or accepted based on the value's comparison to chi-square distribution. The test statistic follows the chi-square distribution with *k-c* degrees of freedom, where k is the total number of observed frequencies and c is the number of parameters being estimated. The null hypothesis is rejected if $\chi^2 > \chi^2_{1-\alpha,k-c}$ where α is the significance level and *k-c* is the degrees of freedom for the chi-square critical value.

Kolmogorov-Smirnov Goodness of Fit Test

Kolmogorov-Smirnov goodness of fit test defines a test statistic, which is essentially a function of the data measuring the distance between the hypothesis and the data. Then it calculates the probability or confidence level of obtaining the data. If the probability is low, then the fit is poor. On the other hand, if the probability is close to 1, then fit is too good to be true and hence a mistake occurred. One advantage of this test is that it is an exact test, so, for example, it does not depend on the sample size as chi-square goodness of fit test does. Two major limitations of this distribution are that it applies to continuous distribution and it is more sensitive at the center of the distribution than at the ends.

Kolmogorov-Smirnov goodness of fit test is based on the Empirical Distribution Function (ECDF). ECDF is a cumulative distribution function that estimates the true cdf of points in the sample. ECDF is a step function that increases by *1/N* at each ordered data point where *N* is the total number of ordered data points. The Kolmogorov-Smirnov test is defined as followed:

Null Hypothesis: Data follows the specified distribution.

Alternate Hypothesis: Data does not follow the specified distribution.

Test Statistic: $$D = \max_{1 \le i \le N} \left(F(Y_i) - \frac{i-1}{N}, \frac{i}{N} - F(Y_i) \right)$$

where F(x) is the theoretical cumulative distribution function of the distribution being tested and Y_i are ordered data.

If *D*, the test statistic, is greater than the critical value, then the hypothesis is rejected at the select critical level, α. The critical values are provided by the software or can be read from a table.

Anderson-Darling Goodness of Fit Test

The Anderson-Darling goodness of fit test compares the fit of an observed cumulative distribution function with that of an expected cumulative distribution function. It gives more weight to the ends than the Kolmogorov-Smirnov test does. Another difference between this and Kolmogorov-Smirnov test is that Anderson-Darling test uses the specific distribution to calculate critical values. This is advantageous because it allows the test to be more sensitive. The Anderson-Darling test is defined as followed:

Null Hypothesis: Data follows the specified distribution.

Alternate Hypothesis: Data does not follow the specified distribution.

Test Statistic:

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\cdot[\ln F(X_i) + \ln(1-F(X_{n-i+1}))]$$

where F(x) is the theoretical cumulative distribution function of the specified distribution and Y_i are ordered data. If A^2, the test statistic, is greater than the critical value, then the hypothesis is rejected at the select critical level, α.

Propositions for Goodness of Fit Tests

The data provided total number of acres burned in each year. A histogram of the dataset was created (Appendix Figure 1). The histogram seemed to resemble the normal distribution. EasyFit software was used to find a distribution that best matched the histogram. Weibull distribution seemed the best distribution to model the number of acres burned. Three types of goodness of fit tests were run to come to this conclusion. Both the Kolmogorov Smirnov test and the Anderson Darling test ranked Weibull Distribution as the best one to fit the histogram. Meanwhile the chi-square goodness of fit test ranked Weibull distribution as the second best with a statistic of 0.6693. It ranked Normal Distribution as the best model. However, the Kolmogorov Smirnov test and the Anderson-Darling test ranked Normal Distribution as the 12th best and 7th best distributions for the model, respectively. According to M. A. Stephens, Anderson Darling test is more robust and a better goodness of fit test than the other two tests[10]. In the end, since Weibull Distribution was ranked first in two tests and second in the chi-square goodness of fit test, it was chosen as the best distribution to model the number of acres burned by fires.

Poisson Distribution and Hurricanes

The Poisson Distribution can be used to model number of events occurring in a time interval. Poisson Distribution has the following probability distribution: $P(X= x) = e^{(-\lambda)}(\lambda)^{x}/x!$, where x is an integer greater than or equal to 0. λ is the shape parameter that is equal to the average number of events in a given interval. This distribution has several statistically special properties including λ being both the mean and variance of the population.

The special property of hurricanes is that they can be modeled after Poisson distribution[8]. To calculate the probability of there being zero hurricane category 1, x=0, plug λ = 0.320 in the distribution function. λ = 0.320 is the average number of category 1 hurricanes occurring in a year. The average was calculated by both the least-squares estimation (LSE) method and the maximum likelihood estimation (MLE) method. MLE method provided an average of 32.07% while the LSE method provided an average of 29.7%. The MLE method is better than LSE because it provides the average that makes the observed data the most probable. Therefore, MLE is less likely to be skewed by outliers.

The resulting probability from the example is 73%, which means the probability of 0 category 1 hurricanes is 73%. Therefore, logically, the probability of 1 or more category 1 hurricanes is 1 - .73 = 27%. This is because probability of an event happening 0 times plus the probability of an event happening 1 or more times should equal 1. This model was used to find the probability of there being at least 1 hurricane from each category. Data was already divided into categories based on the Saffir-Simpson scale. A histogram of the proportion of hurricanes falling in each category is provided in Appendix Figure 6. Rationally, there would be more category 1-3 hurricanes than 4 and 5, because the latter categories require much stronger winds and the probability of that phenomenon occurring is very low.

Tropical Storms

The data also contained tropical storms count and economic damage. Seven categories were assigned to each tropical storm based on the economic damage caused (Appendix Table 4). The cutoffs for the economic damage were assigned so the number of storms in each of the top 5 categories was roughly the same. Categories 4 and 5 had fewer storms, because naturally, the probability of a storm causing such extensive damage while being a tropical storm should be low. Similarly, the tropical storm that caused 5 billion dollar worth of damage should be considered an anomaly. Since tropical storms are smaller hurricanes, they can also be modeled after Poisson Distribution.

Part II: Estimating

Hurricane

An estimate of the damage caused by hurricanes is computed. The Poisson model described above is used. After the probabilities of at least 1 storm were computed for each categorical storm, they were multiplied by the average damage caused by each category. The expected claims were the resultant of the multiplication. The average for hurricanes and Tropical storms were added to determine the total hurricane expected claim.

The tropical storms account for nearly $162 million of damage. This estimate would be much lower, around $60 million, if the data is analyzed without the 2001 tropical storm Allison that caused $5 billion worth of damage. The omitted average resembles the hurricane category 1 expected claim of $53 million. There were 44 category 1 hurricanes in the dataset and the total damage caused by them was $6.85 billion. There were 34 category 2 hurricanes in the dataset and the total damage caused by them was $23.4 billion. The expected claim of $189 million per storm can be expected for category 2. The most number of hurricanes and the most amount of damage were caused by category 3 hurricanes. There were 53 of them in the dataset and the total damage caused by them was $153.6 billion. The expected claim of astounding $821 million per storm can be expected for category 3. There were only 14 category 4 hurricanes and 3 category 5. The total damage was $23.6 billion and $27.9 billion, respectively. The expected claims were $148 and $175 million dollars, respectively. The total damage from hurricanes annually can be expected to be $1.4 billion. The tropical storms add the $162 million to raise the total expected claim to $1.55 billion.

Using the Poisson method, the total damage in 1 year from hurricane is expected to be $1.4 billion. If a straight average is taken without categorizing the storms, the expected damage is $1.37 billion. The sum of each of the average damages in each category should equal the overall average damage. In this case the two averages serve as an indicator of Poisson distribution being a good model for modeling the hurricane data.

Another test to see if Poisson distribution fits the observed data is to verify that average equals variance of the sample. This is because in Poisson distribution, mean equals variance. In this case, the variance of the sample is $1.32 billion. The average is $1.4 billion. These two values indicate Poisson distribution is a good model for the hurricane data.

Fires

The number of acres burned by fires was estimated using the Weibull Distribution. This estimate is compared to an estimate by linear regression.

Comparison to Estimation by Linear Regression

Since Weibull distribution was applied for the first time, another method testing the estimate was needed. Regression seemed the best option to make sure that the estimates of the Weibull distribution fell in the range of the estimate provided by regression modeling.

The data were plotted on a graph as a scatter plot to observe the trends and subsequently model the data. A regression model was developed to model the acres burnt. A linear regression was investigated for the acres burned over the years 1900 to 2008 (Figure 1). The estimated acres burn for 2013 based upon the regression are 4,309,844 ± 2,300,987 acres (95% CI). This value compares favorably 4,486,600 acres, which was estimated based upon Weibull's distribution model.

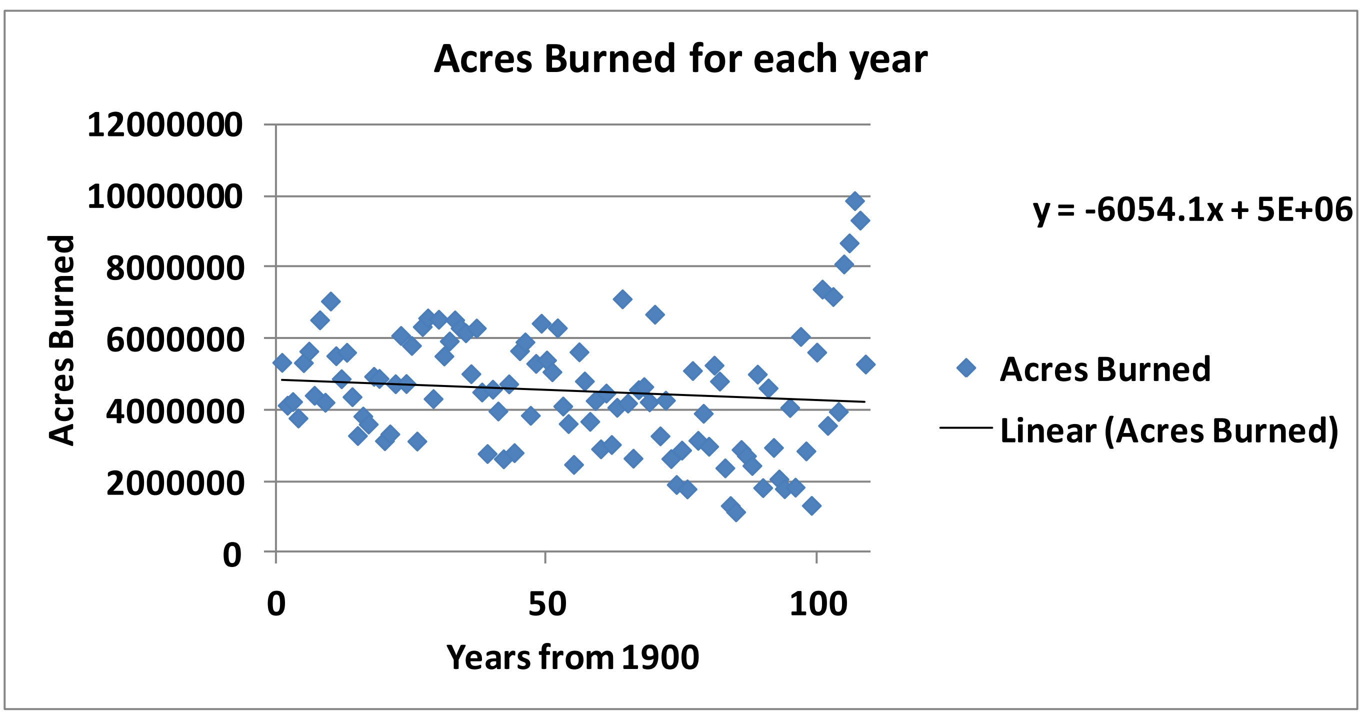


Figure 1: Linear regression for acres burned.

Normality of data was tested as detailed below.

A symmetry plot (Figure 2) gives graph of the distance above the median for the i-th value against the distance below the median for the i-th value. A variable that is symmetric would have points that lie on the diagonal line. The graph shows that this distribution is symmetric.

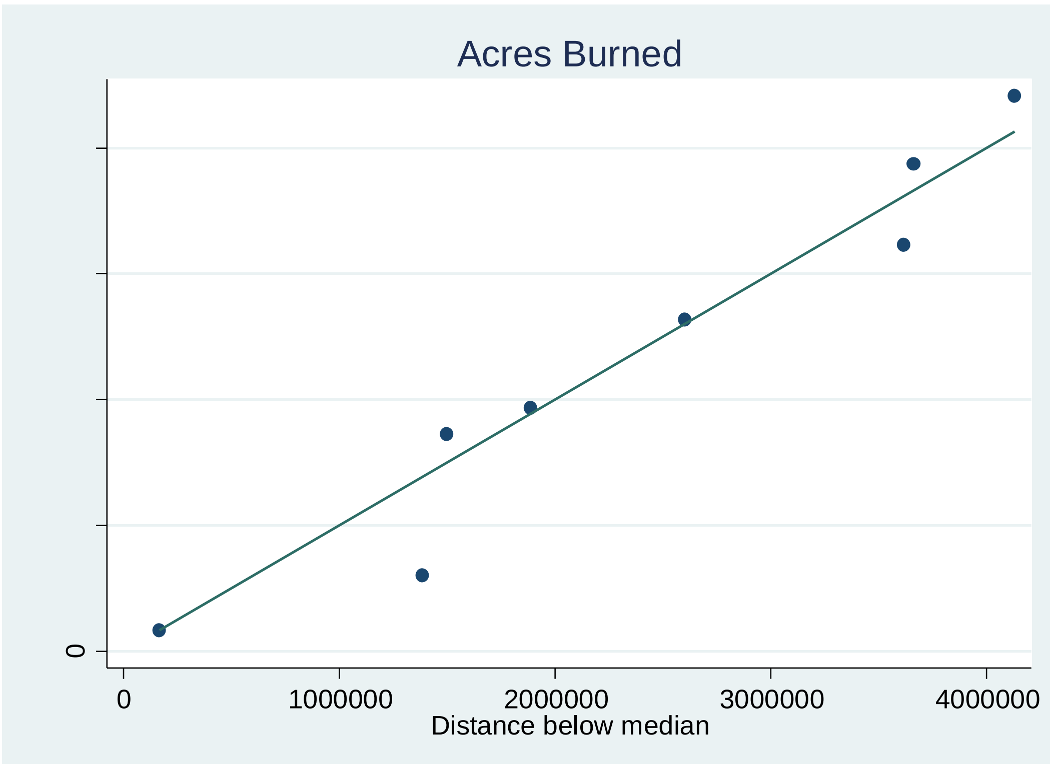


Figure 2: Symmetry plot to test for symmetry of data.

A normal quantile plot (Figure 3) graphs the quantiles of a variable against the quantiles of a normal (Gaussian) distribution. **qnorm** is sensitive to non-normality near the tails, and we do not see considerable deviations from normal, the diagonal line, in the tails.

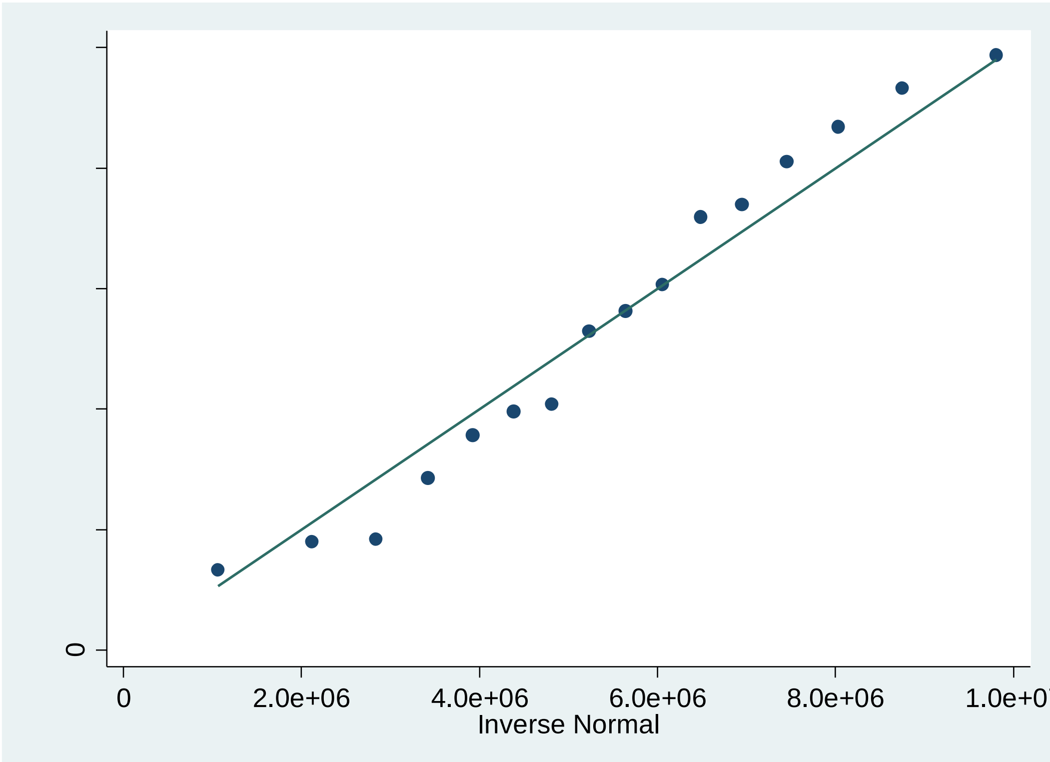


Figure 3: Normal quantile plot to test for normality of data.

The normal probability plot (Figure 4) is also useful for examining the distribution of variables. It is sensitive to deviations from normality nearer to the center of the distribution. There were not any indications of non-normality in acres burned.

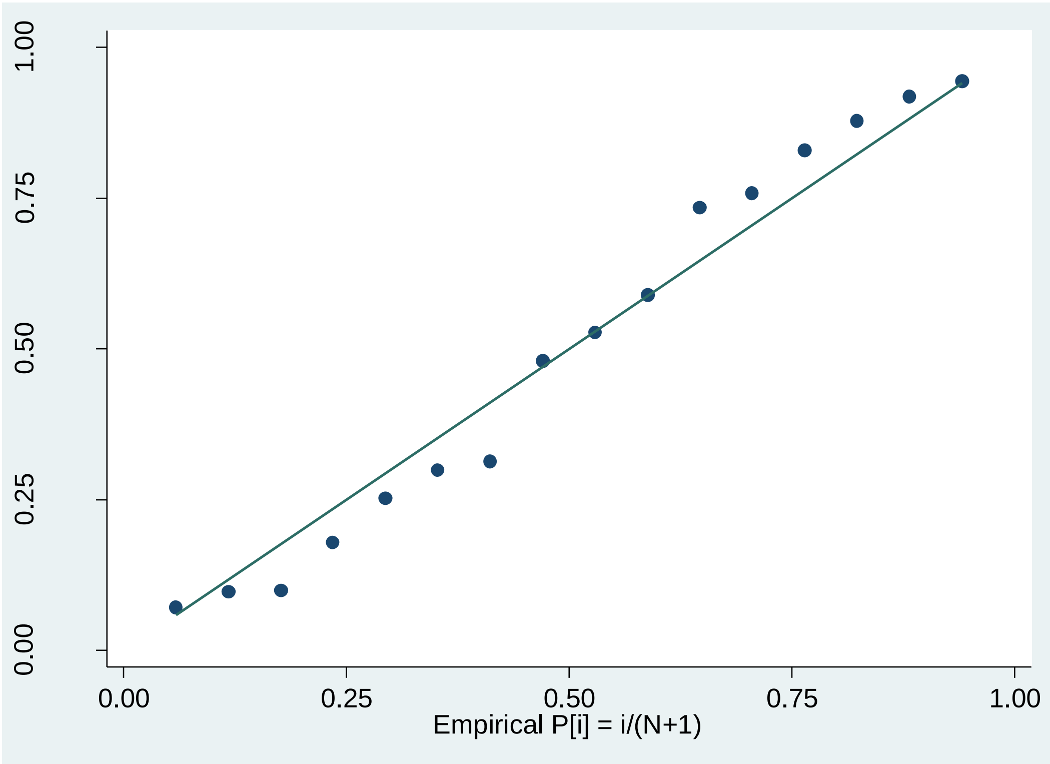


Figure 4 Normal Probability plot (pnorm) deviations for normality check.

Many researchers believe that regression requires normality. This is not the case. Normality of residuals is only required for valid hypothesis testing, that is, the normality assumption assures that the p-values for the t-tests and F-test will be valid. Normality is not required in order to obtain unbiased estimates of the regression coefficients. OLS regression merely requires that the residuals (errors) be identically and independently distributed. If the model is well-fitted, there should be no pattern to the residuals plotted against the fitted values. The residuals graph (using Stata software) depicts no pattern and is shown in Figure 5.

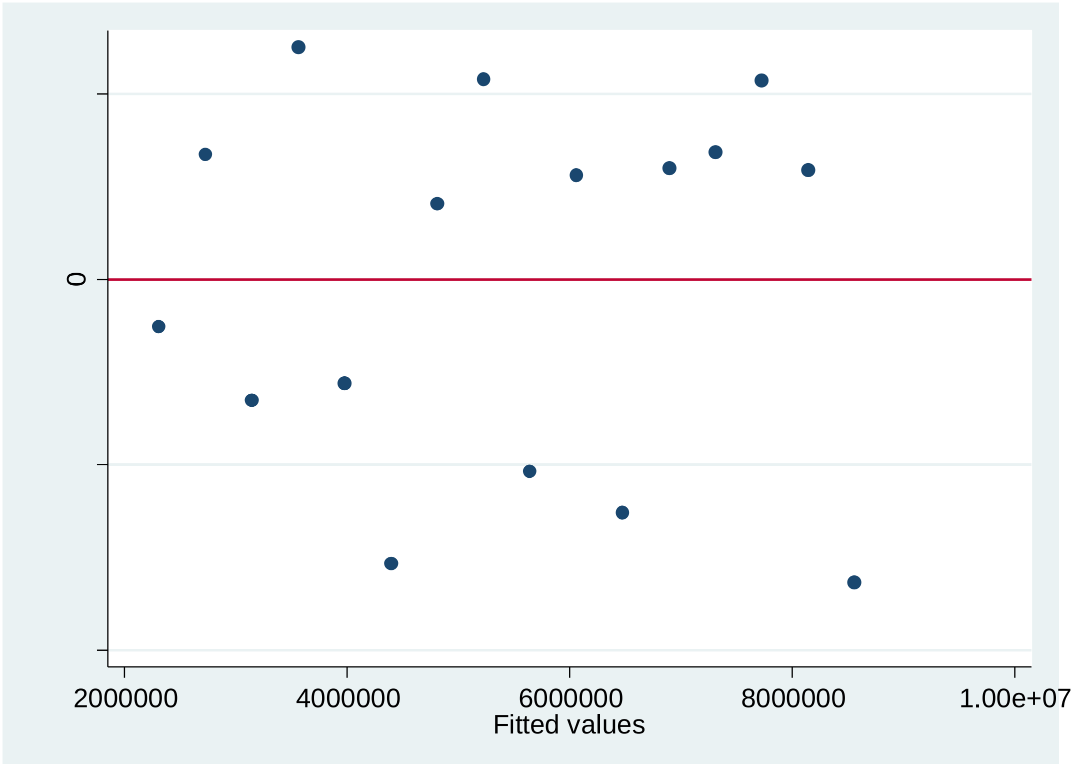


Figure 5: Residuals plotted against the predicted or fitted values of the regression model.

Breusch-Pagan / Cook-Weisberg test[18] for heteroskedasticity was carried out in Stata and the results are as below

Ho: Constant variance

Variables: fitted values of acres burned

chi2(1) = 0.48

Prob > chi2 = 0.4882

This tests the null hypothesis that the variance of the residuals is homogenous. Therefore, if the p-value is very small, we would have to reject the hypothesis and accept the alternative hypothesis that the variance is not homogenous. However, in this case, p = 0.4882 (>0.05) and as such the residuals are homogeneous.

Economic Damage from Fires

To determine the economic damage caused by fires, we need to determine the economic loss per acre of wild fire. The value of land in a national forest where a wildfire can arise does not have a market value but environmental value. We assume each acre has an intrinsic value of $10,000. According to the National Fire Protection Association (NFPA)[22] the total cost of fire in the US was $362.2 billion in 2008. This figure includes property damage, insurance coverage, monetary value equivalent for injured/deceased firefighters and civilians and other economic costs. Since this paper covers only damage from acres burned, economic losses from property damage and other economic losses could be used from the NFPA report. Economic losses from property damage accounted to $20.1 billion, while other economic damage accounted to $44.0 billion.

The acres burned from fires for 2013 are predicted to be 4,486,600 acres. The economic damage from the fires in 2013 can be expected to be 4,486,600 x $10,000 = $44.9 billion.

From the NFPA report, if only property damage estimate is used, the average intrinsic cost per acre would be around $5,000. If both the property damage estimate and other economic damage estimate are used, then the average intrinsic cost per acre would be around $15,000. Therefore an estimate of $10,000 assigned to each acre is reasonable.

Blizzard

A similar regression analysis was carried out for the economic damages for the blizzards from year 1900 to 2006. The data was extrapolated using a linear regression equation fitted to the data. The regression equation modeled is – Economic damage =

Y = -6296074896.755420 +3344175.48 X where X is the year. (Figure 6)

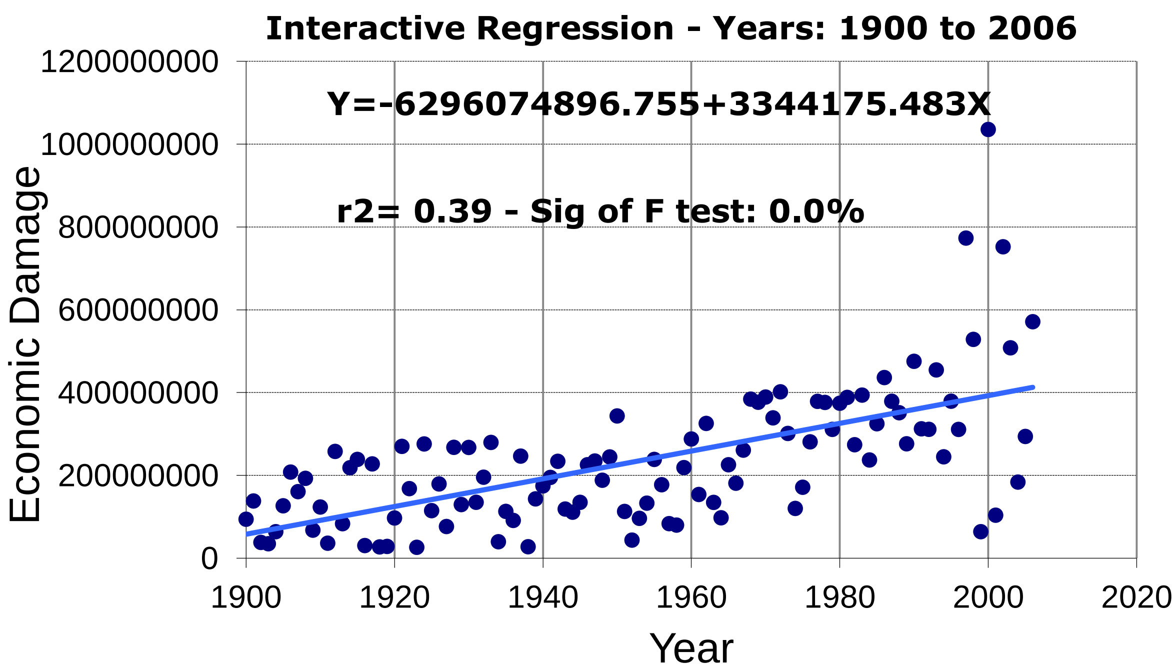


Figure 6. Interactive Regression for Blizzards

For 2013 the economic losses are estimated to be $432,406,175 ± $275,366,172. The residual plot and homogeneity tests confirmed that the assumptions were valid (data or graph not shown).

Poisson Distribution

In addition to the regression, a Poisson model was created for blizzards. The histogram shows the first 3 or 4 bins having the most volume in terms of economic damage and then the economic damage tapers off (Appendix Figure 4). Each bin represents the number of years with at most x hundred million dollars of damage since the last bin. For example, bin 2 represents the proportion of years with economic damage between 100 and 200 million dollars. The fitted Poisson distribution to the histogram is shown in Appendix Figure 5. The model produced a λ value, which is also the mean, of 2.4135. Therefore, on average in a single year, one can expect $241,350,000 worth of economic damage from blizzards. This value is in the range produced by the linear regression.

National Catastrophe Fund Results

The fire can be expected to cause $44.9 billion worth of damage. This implies that the maximum damage caused by fires that the NCIF should expect is 10% or $4.49 billion. The blizzards can be expected to cause damage of $241 million and consequently NCIF can have a reserve of $24.1 million for the purpose.

NCIF can be expected to have a fund of $155 million to cater (10%) for $1.55 billion of expected damages from hurricanes.

The total reserve fund that NCIF has to have is a sum of above three subfunds- $4.49 billion for fire damage, $0.024 billion for blizzards and $0.155 billion for hurricanes. Out of the three catastrophes, in 2013, fires are expected to cause the most economic damage (96.2%) followed by hurricanes and blizzards at 3.3% and 0.5% respectively (Figure 7).

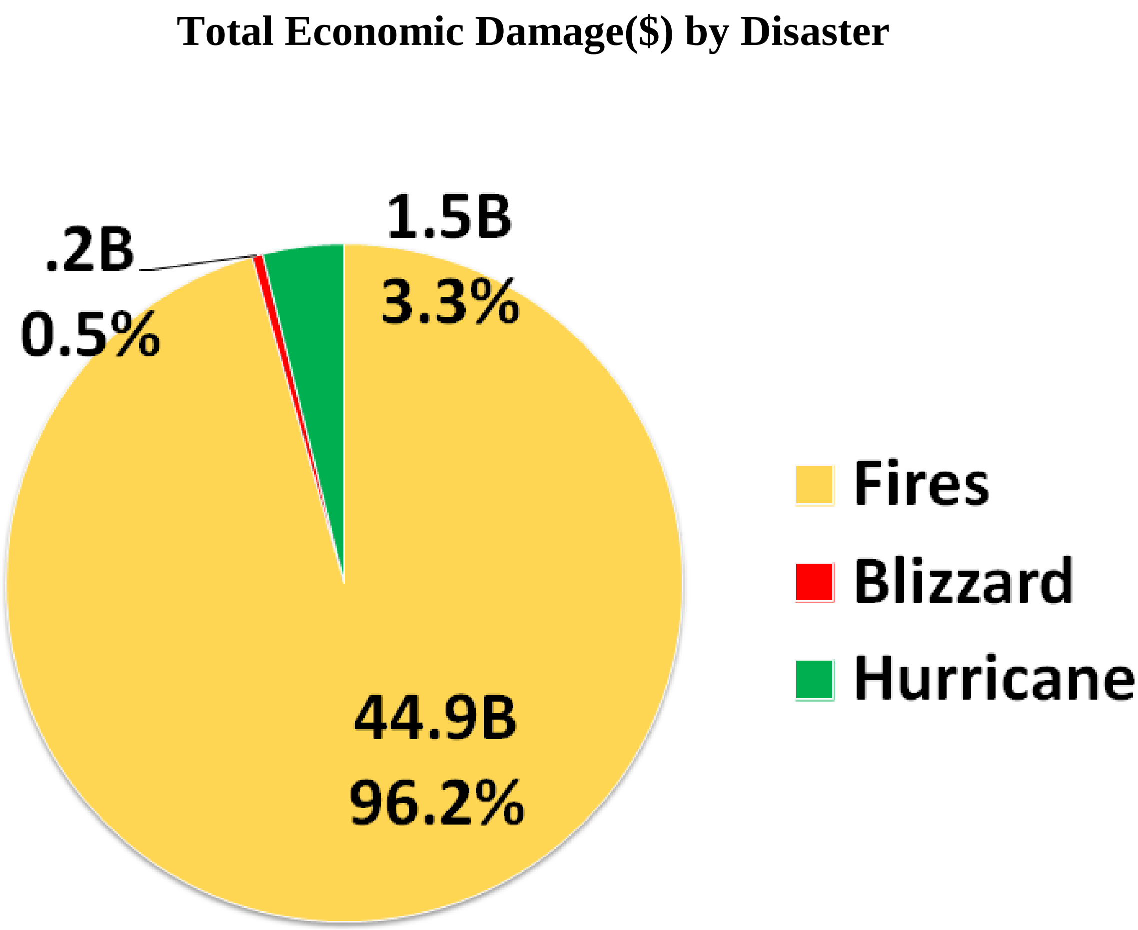


Figure 7. Breakdown of the Total Economic damages caused by the three catastrophes.

Discussion

The National Catastrophe Insurance Fund should expect $4.67 billion to be expected claims for the upcoming year. This information is based on the Poisson distribution model for hurricanes and blizzards and Weibull Distribution for Fires. The assumptions made for fires was economic loss per acre burnt for an intrinsic value of $10,000 based on the NFPA report. There is a lot of variability in blizzard damage in the last few years. This effect could be due to natural effects, such as global warming, or a data inconsistency. With increased housing today compared to the early 20th century, the definition of blizzard and its differentiation from winter storm may have changed. However, the economic damage in recent years has been variable, depending upon if the blizzard of same intensity struck highly populated areas or hardly habituated areas. The limitations for both fires and blizzards include inability of a proper model to accurately predict either of the catastrophes.

The hurricanes and tropical storms can also be predicted using the Poisson distribution. Most of the damage of the hurricanes comes from category 3 hurricanes. This expectation is augmented by the fact that most hurricanes are also category 3. However, this is just a coincidence in the data as all storms below and equal to category 3 are common, because attaining wind speeds equal to higher category storms is much more difficult[6]. The claim of hurricanes expects to be near $1.55 billion. The tropical storms account for $162 million of it. However, the number for tropical storms would be close to $60 million, if the 2001 storm Allison had not caused $5 billion of damage. This storm by itself skewed the probability of a devastating tropical storm and increased the average of a tropical storm by $86 million.

Conclusion

The NCIF with according to 10% reserve should have enough funds to allocate $4.67 billion. The models are just a prediction for year 2013 based on averages and trends from the last 100 years or so. As always more data and more precise models can further accurately predict the allocation of funds needed to cover economic damages. The actual claims could be much higher or much lower and would depend on the extent of a catastrophe. The extent does not always depend on the force of the natural disaster. It can, for example, depend on the preparedness of the government to battle a disaster[20] or even on the conditions of the structures impacted by the disaster.

References

1. The total cost of Fire in United States. http://vmw.nfpa.org/assets/files/PDF/ totalcostsum.pdf, Accessed 10/05/2011

2. Weisman, Brent. Hurricanes in Florida. http://fcit.usfeduifloridailessons/hurricane/ hurricane.htm. Accessed 10/02/2011

3. Rita F. Fahy, Paul R. LeBlanc and Joseph L. Molis. Firefighter Fatalities in the United States —2010. by NFPA, Quincy, MA. Accessed 10/05/2011

4. Hamid, Shahid. Public Hurricane Loss Projection Model. Florida International University. Accessed October 2011

5 Strengthening Disaster Recovery for the Nation. FEMA. http:// www.disasterrecoveryworking,group.govioutreach/regionIX-pacific-rim-vtc-121409.pdf Accessed October 2011.

6. Anthnes, Richard A. Data Assimilation and Initialization of Hurricane Prediction Models. Journal of Atmospheric Sciences, vol. 31, Issue 3, pp.702-719

7. Wright, Henry A. and Arthur W. Bailey Fire ecology, United States and southern Canada. Pp. xxi + 501. John Wiley and Sons.

8. Katz, Richard, W. Stochastic Modeling of Hurricane Damage. Journal of American Meteorology. Volume 41, Pg. 754-762.

9. Watkins, Joseph. Topic 13: Unbiased Estimation. Department of Mathematics - University of Arizona. November 3, 2009. http://math.arizona.edu/~jwatkins/m-unbiased.pdf. **Accessed 3/30/2013.**

10. Stephens, M. A. (1974). "EDF Statistics for Goodness of Fit and Some Comparisons". ***Journal of the American Statistical Association*** **(American Statistical Association) 69 (347): 730–737.**

11. Weibull, W. (1951), "A statistical distribution function of wide applicability", ***J. Appl. Mech.-Trans. ASME*** **18 (3): 293–297.**

12. *Engineering Statistics Handbook*. 1.3.5 Quantitative Techniques. National Institute of Science and Technology. http://www.itl.nist.gov/div898/handbook/eda/section3/eda35.htm. **Accessed 3/30/2013.**

13. H.R. 2555: Homeowners' Defense Act of 2010. Rep. Ron Klein. http://www.govtrack.us/congress/bills/111/hr2555. Accessed 3/30/13.

14. SAS. Business Analytics and Business Intelligence Software. SAS v9.3. SAS Institute.

15. Stata. Data Analysis and Statistical Software. Stata v12. StataCorp LP.

16. Dahl, J. et al. Maximum likelihood estimation of Gaussian graphical models:

Numerical implementation and topology selection. UCLA Preprint, 2005. http://www.ee.ucla.edu/~vandenbe/publications/covsel1.pdf. **Accessed 3/30/13.**

17. Probability-Probability (P-P) Plot. York University. http://wiki.math.yorku.ca/images/0/08/Statistics_pp-qq-1.pdf. **Accessed 3/30/13.**

18. Prof. Sharyn O'Halloran. Sustainable Development U9611 Econometrics II. Columbia University. http://www.columbia.edu/~so33/SusDev/Lecture_6.pdf. **Accessed 3/30/13.**

19. Probability Distribution of Weibull Distribution dependent on Shape Parameter. Easy Fit Software - Microsoft Excel. Accessed 3/30/13.

20. *Food and Water in an Emergency*. FEMA. August 2004. http://www.fema.gov/pdf/library/f&web.pdf. **Accessed 3/30/13.**

21. Crépon, Bruno. Maximum Likelihood Estimation. Example : Poisson Distribution. http://catalogue.polytechnique.fr/site.php?id=335&fileid=4795. **Accessed 4/18/13.**

22. Hall Jr., John R. The Total Cost of Fire in the United States. National Fire Protection Association Fire Analysis and Research Division - February 2011. http://www.nfpa.org/assets/files/pdf/totalcostsum.pdf. **Accessed 4/18/13.**

Appendix

Table 1. Sample data for hurricanes.

Year	State	Storm Category	Economic Damages ($)
1926	FL,AL	4-3	105,000,000
1928	FL	4	25,000,000
1929	TX	1	700,000
1929	FL	3	300,000
1932	TX	4	7,500,000
1933	VA	2	17,000,000
1933	TX	3	12,000,000
1933	FL	3	1,000,000
1933	NC	TS	1,000,000
1934	SC	3	200,000
1934	LA	3	2,600,000
1934	TX	2	1,500,000
1935	FL	5	6,000,000

Table 2. Sample data for fires.

Year	State	Count	Acres Burned
1900	NM, WA, AK, NC	101066	5339650.34
1901	WA, CA, AK	132504	4137375.58
1902	CA, NM, NV, OR	106257	4240369.70
1903	OR, CA, MT	59994	3779034.13
1904	OR, CA, ID	78740	5331782.23
1905	OR, AK, CA, AR	111134	5653690.84
1906	CA, MT, ND, NM	108965	4415242.34
1907	NM, OR, UT, CA, AK	117065	6532465.37
1908	CA, WA, NM	110395	4220611.48
1909	NM, UT, NC, AR	75051	7059476.68
1910	WA, OR, SD	96985	5525697.99
1911	AK, OR, CA, MT	70492	4880255.87
1912	UT, CO, NM, AR, CA	72943	5621221.97

Table 3. Sample data for blizzards.

Year	State	Count	Economic Damages ($)
1900	PA, NY, NJ	3066	$93,986,000
1901	NY, MN, ME	3028	$138,061,000
1902	ME, NJ, PA, NY	2981	$37,990,000
1903	MI, PA, CT, MN	3187	$34,845,000
1904	NY, VT, OH, NC	3324	$63,946,000
1905	PA, NY, MN	2934	$126,481,000
1906	ME, NY, CT, PA, MN	2820	$207,842,000
1907	MI, NY, MN, NJ	2964	$160,657,000
1908	PA, NJ, MN, MI, OH	2990	$192,712,000
1909	MI, NY, TN	2598	$67,587,000
1910	CT, MN, NY	2895	$123,496,000
1911	PA, TN, NY, IL, MN	2823	$36,137,000
1912	MI, NY	2767	$257,780,000

Table 4. Cutoffs and Count for Tropical Storms

TS category	Cutoff	Count
1	<2 million	12
2	2-5 million	13
3	6-24 million	12
4	25-99 million	10
5	>100 million	10
6	>1 billion	1
7	STS	1

Figure 1. Histogram of Number of Acres burned.

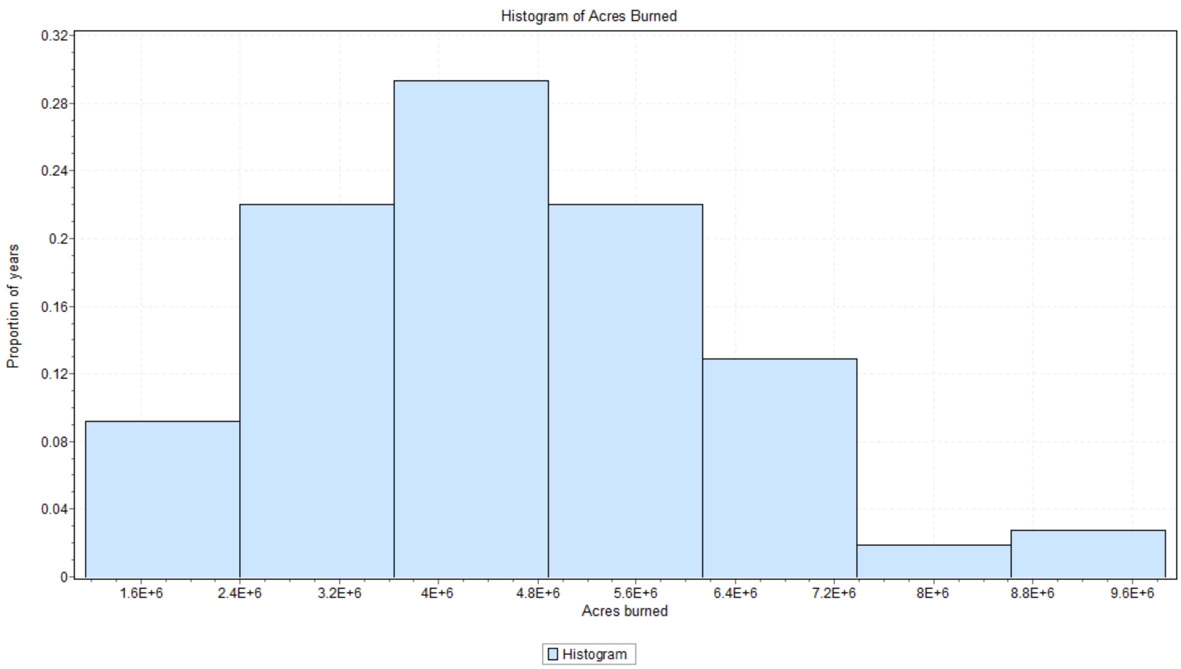


Figure 2. Weibull distribution for acres burned by fires.

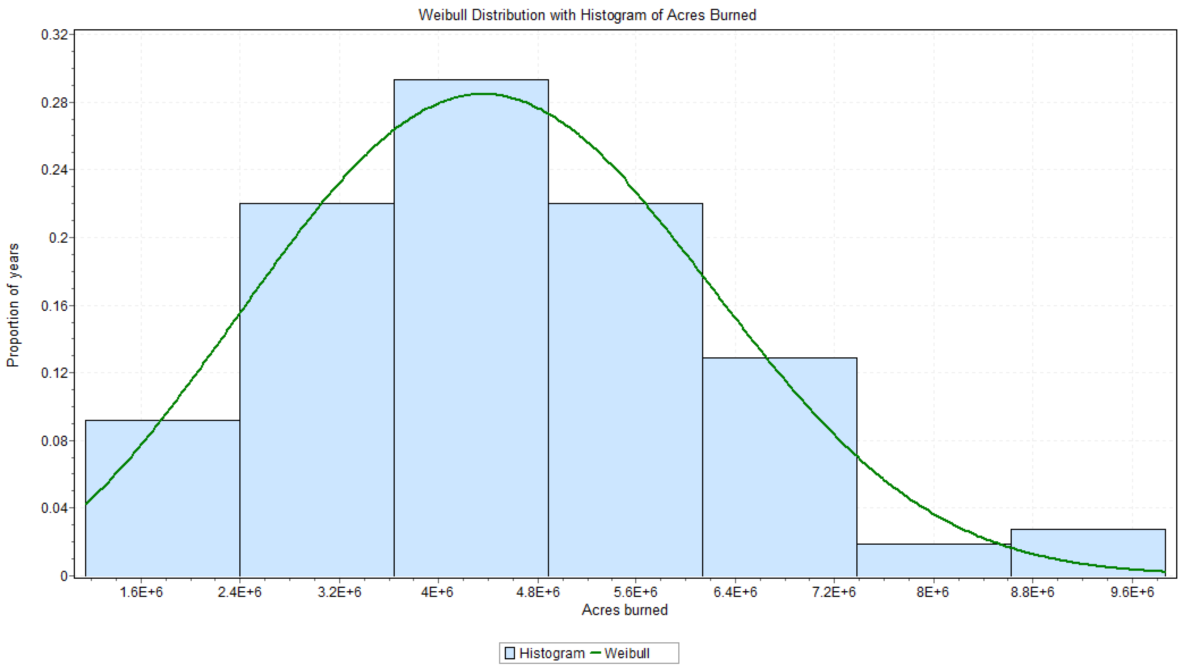


Figure 3. Normality test for Weibull Distribution.

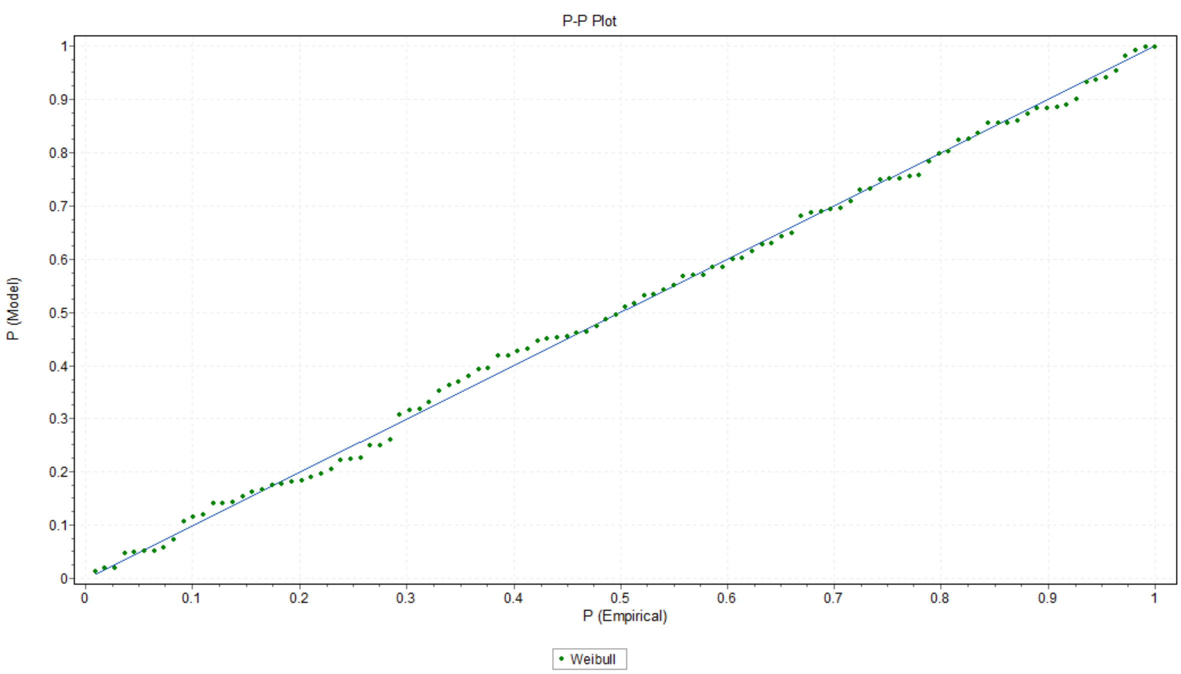


Figure 4. Histogram for Economic Damages caused by Blizzards

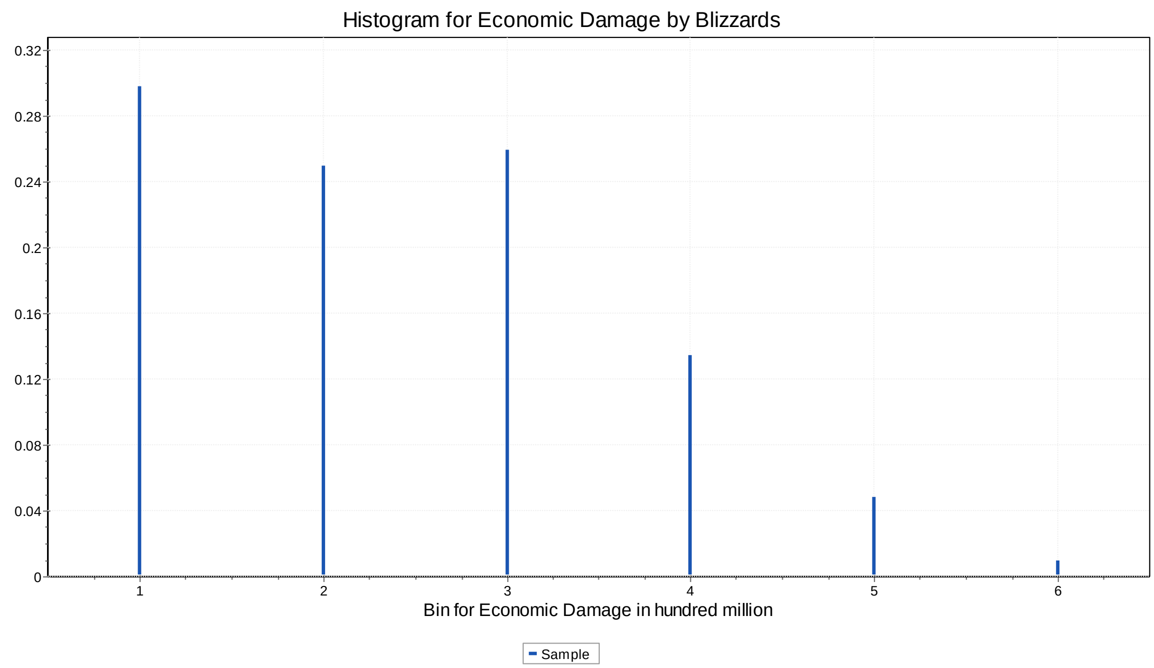


Figure 5. Poisson Distribution with Histogram for Economic Damages caused by Blizzards

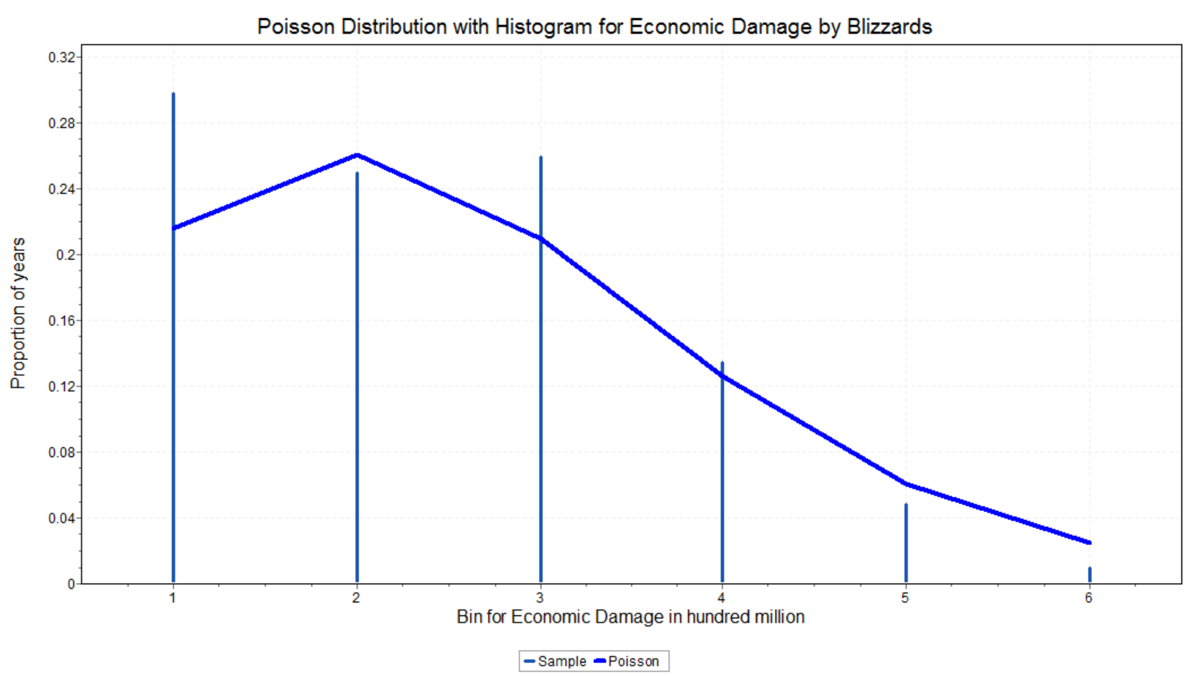


Figure 6. Histogram of the Proportion of Hurricanes in each Category.