Risk Assessment of Nuclear Power

Executive Summary

This project examines the feasibility and safety of nuclear energy as a power source. A list of all nuclear power plant accidents in history is examined, and the frequency of accidents in any given year is modeled with a negative binomial distribution. Not only do the results show that nuclear power plant accidents are rare, they also show that the majority of all accidents have very low costs. This confirms the hypothesis that the dangers of nuclear power are exaggerated by a few case examples (such as Chernobyl.)

Background & Motivation

The human race is accelerating the onset of diseases, natural disasters, and biodiversity loss to the point where it is leading to our own extinction, making environmental degradation one of humanity's most significant problems. 99.9% of the academic research papers published over the last two years agree that global warming is real and caused by humans[1].

Greenhouse gasses trap infrared radiation (heat) in the atmosphere, which is an essential process because it creates necessary conditions for life on Earth. Naturally the influx of greenhouse gasses is at an equilibrium level with sequestration, photosynthesis, and other processes that take greenhouse gases out of the atmosphere; however, when humans burn oil, coal, and gas for energy, unusually high amounts of greenhouse gasses are emitted into the atmosphere to the point where the equilibrium is upset and the earth's

---

[1] http://www.msnbc.com/msnbc/how-climate-change-deniers-got-it-very-wrong
[2] https://medium.com/matter/it-s-not-climate-change-it-s-everything-change-

temperature keeps rising. Substantial scientific research has proven that the speed and magnitude of this warming effect is so drastic that current levels of fossil fuel consumption are simply too high to sustain our generation.[2]

This begs the obvious question: what can we do about it? Within the framework of the current world population it is not an option to simply reduce energy consumption because all of the global economy (transportation, manufacturing, cities, etc.) is completely dependent on electricity. There are currently 3 types of energy: fossil fuels (coal, oil, natural gas), renewables (solar, hydro, wind, etc.), and nuclear. It is clear that fossil fuels are completely unsustainable (they will run out soon if we keep using them) and also incredibly dangerous (they cause global warming.) The majority of the opposition to fossil fuel consumption is coming from hard-core "only-renewables" environmental activists, but unfortunately implementation of renewables requires astronomical upfront costs (i.e. a solar panel or wind turbine requires 10+ years to be more financially attractive than fossil fuels.)

Nuclear power fills a unique niche because it does not emit greenhouse gases. Moreover, nuclear power plants become profitable at a much faster rate than any other renewable energy sources, which is why they account for the majority of emissions-free power production.[3] The only true downside to nuclear energy is the low-probability high-magnitude risks it presents in the circumstances of an accident, which is the major source of opposition to it. My hypothesis is that risks associated with nuclear power are greatly outweighed by the benefits it provides; I hope to show this examining the

---

[2] https://medium.com/matter/it-s-not-climate-change-it-s-everything-change-8fd9aa671804

[3] http://www.nei.org/CorporateSite/media/Images/Infographics/Sources-of-Emission-free-2014-01.png?width=1500&height=1125&ext=.png

likelihood and severity of all the past nuclear power plant accidents. I will do this analysis by examining a list of every nuclear power accident in history including it's year, location, and severity.[4]

The Data

A quick glance at the list indicates that there have only been 33 accidents in the 66 years since 1950. Moreover, accidents of higher severity are way less likely than accidents of low severity. These observations prompted me to examine just how unlikely it is for a devastating nuclear accident to occur.

Year	Incident	INES level	Country
2011	**Fukushima**	5	**Japan**
2011	**Onagawa**	1	**Japan**
2006	Fleurus	4	Belgium
2006	Forsmark	2	Sweden
2006	Erwin	1	US
2005	Sellafield	3	UK
2005	Atucha	2	Argentina
2005	Braidwood	1	US
2003	Paks	3	Hungary
1999	Tokaimura	4	Japan
1999	Yanangio	3	Peru
1999	Ikitelli	3	Turkey
1999	Ishikawa	2	Japan
1993	Tomsk	4	Russia
1993	Cadarache	2	France
1989	Vandellos	3	Spain
1989	Greifswald	1	Germany
1986	Chernobyl	7	Ukraine (USSR)
1986	Hamm-Uentrop	1	Germany
1981	Tsuraga	2	Japan
1980	Saint Laurent des Eaux	4	France
1979	Three Mile Island	5	US
1977	Jaslovské Bohunice	4	Czechoslovakia
1969	Lucens	1	Switzerland
1967	Chapelcross	1	UK
1966	Monroe	1	US
1964	Charlestown	1	US
1959	Santa Susana Field Laboratory	1	US
1958	Chalk River	1	Canada
1958	**Vinča**	1	**Yugoslavia**
1957	Kyshtym	6	**Russia**
1957	Windscale Pile	5	UK
1952	Chalk River	5	Canada

Table 1: A comprehensive list of all nuclear power accidents to date.

[4] http://www.theguardian.com/news/datablog/2011/mar/14/nuclear-power-plant-accidents-list-rank#data

The thing that got me excited about this dataset—that there are *only* 33 accidents—is also one of the concerns I have going forward with it. 33 is a substantial enough sample to built a count model with, but I am limited in the number of different subsamples I can analyze (i.e. covariates like countries, severities, etc.)

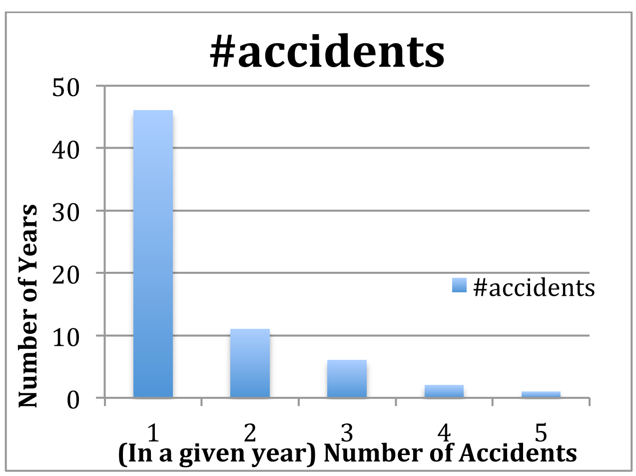
My first step in preparing the data for a count model was to convert the above table into a frequency table such that each year between 1950 and 2015 had the number of accidents associated with it. Table 2 and Graph 1 show the distribution of the raw data.

year	#accidents
1950	0
1951	0
1952	1
1953	0
1954	0
1955	0
1956	0
1957	2
1958	2
1959	1

Table 2 (excerpt): the number accidents in each year

Mean: .5

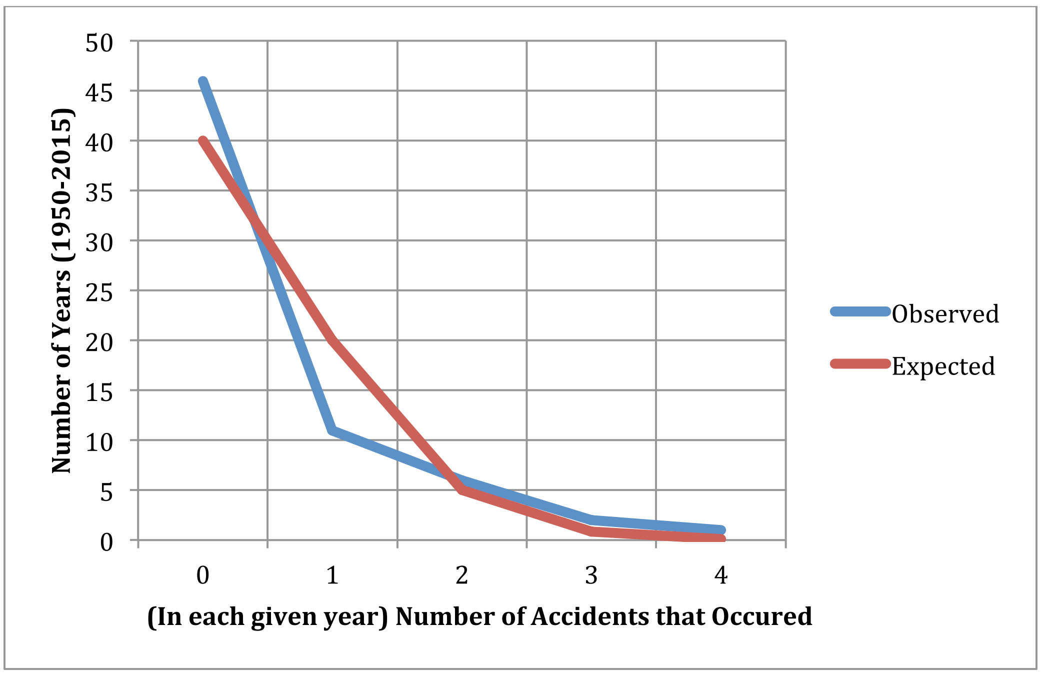
Variance: .81



Graph 1: Histogram of Accidents in Each Year (raw data)

Data Analysis

Upon converting this table into a format that I can build models from, the most urgent next step was to check whether or not this data was homogeneous. If the odds of nuclear disasters were the same year-to-year then there would not be any source of heterogeneity, meaning that the likelihood of an accident is just determined by another "spin of the Poisson wheel." To test this, I built a Poisson distribution (which assumed homogeneity among the different years.)



Graph 2: Empirical and Poisson Distributions of the Number of Nuclear Accidents in a Given Year

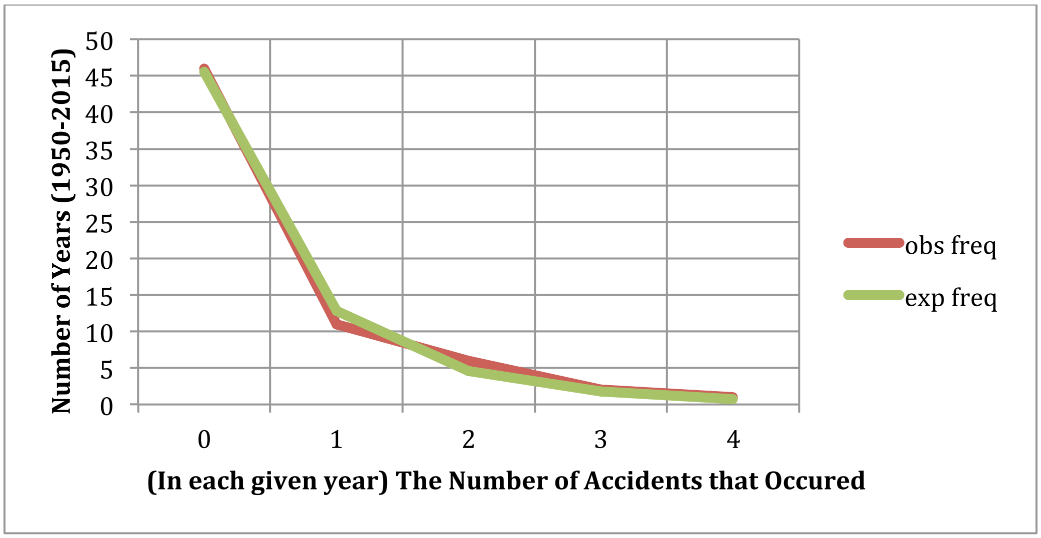
Test Statistics					
lambda	0.5				
sumLL	-66.79				
	Chi Sq Test:	#accidents	obs freq	exp freq	chiSq
		0	46	40.03	0.89
		1	11	20.02	4.06
		2	6	5	0.2
		3	2	0.83	1.65
		4	1	0.1	8.1
				sum:	14.9
				df:	3
				p-value:	0.0019

Table 3: Parameter, Log Likelihood, and Chi-squared values for Poisson model

The lambda of the distribution is .5, which means that the mean and the variance are both .5; since a disaster happened 33 times in 66 years there is, on average, .5 disasters per year. The fit of the Poisson model doesn't appear to be impressive, which is validated by the fact that the p value for the chi squared test is .002, meaning that it is incredibly unlikely that the Poisson model is an accurate representation of the number of nuclear accidents in any given year. (It is worthwhile to note that not each cell in the chi-squared analysis has a frequency equal to or greater than 5, so for both the Poisson and the NBD it is important to take the chi-squared with a grain of salt. That being said, in both cases the p values are consistent with how well the model fits the data in the histogram.)

This poor fit should be unsurprising because different years have different sets of variables that change the likelihood of a nuclear disaster. Laws, standards, number of power plants, whether there is a war, and energy demand are all factors that change on a year-to-year basis and significantly impact the likelihood of a nuclear power plant accident. This is why the Poisson model didn't fit well; the likelihood of an accident next year isn't just a re-spin of this year's Poisson wheel because each year has it's own underlying propensity to have a nuclear accident.

Because there is no doubt that the distinct years are the source of heterogeneity, the next logical step was to build an NBD model that accounts for the heterogeneity. My guesstimate was that the mean of the distribution would be the same, but the variance would be slightly different since it's no longer bounded to be equal to the mean (as in the Poisson case.)



Graph 3: Empirical and Negative Binomial Distributions of the Number of Nuclear Accidents in a Given Year

Before even examining any of the parameter estimates or goodness-of-fit tests, the NBD clearly did a much better job than the Poisson at estimating this behavior. This is even more evidence for the fact that each year must have its own underlying propensity of having a nuclear accident.

r	0.646		variance	(r/alpha) + (r/alpha^2)	0.89
alpha	1.293		mean=r/alpha		0.5
			sumLL		-62.77
Chi Sq Test:	#accidents	obs freq	exp freq	chiSq	
	0	46	45.573	0.004	
	1	11	12.847	0.266	
	2	6	4.613	0.417	
	3	2	1.775	0.029	
	4	1	0.706	0.122	
			sum:	0.838	
			df:	2	
			p-value:	0.6577	

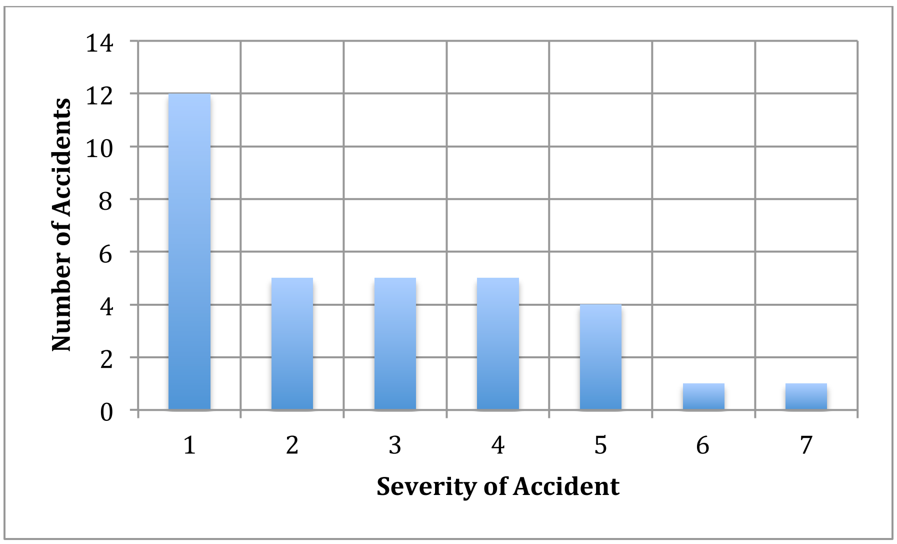
Table 4: Parameter, Log Likelihood, and Chi-squared values for NBD model

The mean of this distribution is also .5, which (again) makes sense because there are 33 accidents in 66 years. Unlike the Poisson, the NBD's variance differs from it's mean. The higher variance in the NBD may explain why the model fits a lot better. The r-value being less than 1 explains why the peak of the distribution is located at zero (i.e. most years don't have any nuclear power plant accidents.)

The better fit is also evidenced by the higher log likelihood as well as much better chi-squared p value. Since the p-value of the chi-squared test is greater than .2, we can confidently say that the expected NBD and empirical histogram frequencies of nuclear power plant accidents are not significantly different.

To confirm this result, I also used the Method of Moments as well as the Means and Zeros method of finding parameters and evaluating fit. Both methods resulted in almost identical values for r, alpha, variance, and mean when compared to the MLE method for the NBD. The Means and Zeros method produced an NBD that fit really well on the observed data (with a p value way above .2); however, the Method of Moments was somewhat less accurate (with a p value of .15) because it's P(0) count was 10 fewer than the observed data. Overall, these two procedures uphold the validity of the MLE estimation that was conducted in order to fit the NBD model.

If there were an accident the scale of Chernobyl an average of .5 times any given years, then I would concede that nuclear power might be a little too risky. It is difficult to get a complete risk analysis from the NBD above because it doesn't account for severity of accidents, which prompted me to examine just how deadly each of the accidents was. Combining our previous analysis with an examination of the histogram of the severities of the accidents provides more insight:



Graph 4: Frequency of Nuclear Power Plant Accidents by Severity

It is immediately evident that the higher severity accidents are a lot less likely to occur than less severe accidents, but this still isn't that informative without knowing exactly what the severities correspond to. The severities are given by the International Atomic Energy Agency's scale INES, which categorizes occurrences of severity 1-3 as "incidents" (i.e. power plants had to pause operation, but no damage/fatalities occurred) and accidents of severity 4-7 as "accidents".[5]

Of the occurrences, 22 are "incidents" and 11 are "accidents". Of the eleven accidents, only 2 could be attributed to any deaths: Chernobyl (31 died in the accident and 19 died over the subsequent 20 years) and Tokaimura (nobody died in the accident and 2 died from radiation over time.)

Accidents	Death Toll	Sources	
Fukushima	0	http://www.unscear.org/unscear/en/fukushima.html	
Fleurus	0		
Tokaimura	2	http://world-nuclear.org/information-library/safety-and-security/safety-of-plants/tokaimura-criticality-accident.aspx	
Tomsk	0		
Chernobyl	50	http://www.nuclearfaq.ca/cnf_sectionD.htm#x	"The immediate significant health effect of the accident was the death toll (mostly due to high radiation exposure) among the workers at the plant itself. Of this group, 31 died either at shortly after the accident, and 19 died subsequently over the next two decades (some not attributable to the accident however)."
Saint Laurent des	0		
Three Mile Islan	**0**	**http://www.ans.org/pi/resources/sptopics/tmi/faq.php**	"No one died as a result of the TMI-2 accident. The accident caused concerns about the possibility of radiation-induced health effects, principally cancer, in the area surrounding the plant. Because of those concerns, the Pennsylvania Department of Health maintained for 18 years a registry of more than 30,000 people who lived within five miles of Three Mile Island at the time of the accident. The state's registry was discontinued in June 1997 without any evidence of unusual health trends."
Jaslovské Bohunic	0		
Kyshtym	0	http://www.mdpi.com/1660-4601/6/1/174	"The study showed that there was no increased number of cancer deaths in the progeny of exposed persons compared to the persons in the control group."
Windscale Pile	0		
Chalk River	0	http://www.nuclearfaq.ca/cnf_sectionD.htm#x	

Table 5: Nuclear Power Plant Accidents and Their Death Tolls

Taking this as well as the NBD into account, we see that in any given year the likelihood of a real accident (not just an incident) is only 1/6. In only 2 of the 11 real accidents were scientists able to identify a death toll, which brings the likelihood of any given year having a deadly nuclear accident down to (1/6)(2/11) = 2/66 = 1/33.

Conclusion and Discussion

When compared to the hundreds of thousands of lives that are lost every year due to starvation, natural disasters, and other climate-changed induced problems[6], it is pretty evident that if we quintupled our nuclear production (the increase needed to replace fossil

[5] http://www-ns.iaea.org/tech-areas/emergency/ines.asp
[6] http://www.gci.org.uk/Documents/MUSE.pdf search for "300,000"

fuels) the expected amount of lives lost would be undeniably lower than the damage we incur by burning fossil fuels. A greater amount of nuclear power plants would also give us more data to work with in terms of examining and minimizing the risks of nuclear power.

This result prompted me to get an even more precise window on the risk of nuclear disasters, which is why I tried to build models for different subgroups of this dataset including divisions into country, continent, and decade. The goal of this analysis was to identify whether certain factors (such as location or year) had high correlation with the likelihood of an accident. Unfortunately, as I had been concerned (in the beginning of the paper) the size of the dataset was already pretty small, so dividing it into subgroups lead to distributions that couldn't accurately be modeled by an NBD because of the inherent randomness of a small dataset. I also tried to do some analysis on how the number of nuclear power plants changes in each country over time, but unfortunately there is only data on how many power plants each country currently has, which isn't enough to produce a narrower window on the probability of an accident.

As a statistician, I wish that my original dataset were larger so that I could break the data into categories for countries, decades, etc. in order to build sub-models that exposed more of the covariates behind nuclear power plant accidents. Those models would have been very useful in not only furthering the proof of how exaggerated the dangers of nuclear power are, but they would have also provided useful insights into where and when it is safest to produce nuclear power. (As someone concerned about the state of the environment, I don't wish that the original dataset were larger because that would mean more accidents among nuclear power plants, which would indicate that they are in fact unsafe to use.)

An important next step is gathering as much information as possible from this moment onward about nuclear power plants. Keeping track of the number of power plants in the world over time will help improve the NBD analysis by providing the necessary information needed to account for the fact that different years have varying numbers of power plants.

This analysis indicates that their dangers are over-exaggerated, or at the very least a lot less menacing than those of fossil-fuel based power plants. A good move would be to increase the number of nuclear power plants while equipping both the new and old ones with precise tools that would measure any problems as they occur so that statisticians can achieve greater certainty when analyzing the risks associated with nuclear power.