Charles Breitenfeldt
Statistics
Final Project

Severe Weather Deaths in the United States

Introduction:

Every year hundreds of people are killed in the United States by weather related events. Therefore we would want to find out how many people die each year and what kind of weather killed them. There are seven different categories that the Storm Prediction Center in Norman, Oklahoma uses for classifications of weather related deaths and each one of those seven categories have subcategories, totaling 26. Out of those 26 subcategories come out what is called "the big 4" of weather related events. Those "big 4" include tornadoes, hurricanes, floods, and lightning; which is responsible for approximately 41% of weather related deaths since 1995.

Being a meteorology student I have always loved weather, especially severe weather. It is the most interesting part of the meteorology field and where all of the major research dollars are spent. There are millions of dollars spent each year on research for the 'big 4" in trying to figure out ways of how these complex events happen, what we can do to better predict when and where they will happen, and also how to design buildings to better protect people from these devastating events. And with such a high percentage of deaths by the "big 4" I was interested in how many people that actually translated into instead of just a percentage. I was also interested to see if it is worth spending all of this money on researching these severe weather events, money that comes from the government as a part of our taxes. I thought that many people would want to know if the money being spent on this research was going to pay off for them in the long run.

With this in mind I decided to research how many people were killed each year by weather related events, particularly to answer the question "at the 1% significance level whether

at least 300 are killed each year by the big 4". I have a feeling that we will find out that there are more than 300 people killed each year by the big 4. In this case the individual would be a year in the United States. The variable would be the number of human beings killed in the United States by the big 4 each year; and the parameter would be all of the years in the United States. After I determined these things I set out to do my research and find out how many people actually died from the big 4.

Methods:

In order to conduct this research I headed to the National Weather Service main website and started so search for archived data. The National Weather Service has archived data going all the way back to their inception on February 9, 1870. The trick was to find the data I was looking for, people killed by the big 4. I found what I was looking for in the Office of Climate, Water, and Weather Services. They had on link on their page for weather fatalities. On this link they had weather related fatalities for all of the 26 subcategories dating back to 1940 and running all the way up to 2007. I took this data and put it into an excel spreadsheet so that I could better look and it and then added a totals column to see what the actual totals were for the last 68 years. I could not obtain information for deaths related to the big 4 before 1940, the only records on file dating back that far were things like temperature, winds, dew points, and what was going on. The National Weather Service was not so worried about weather related deaths until around 1940. It should be noted that there was no randomization used in the collection of this data. Instead I simply took the data from the NWS and used all 68 years of it, to increase my sample size and meet the assumptions for a 1-sample t-test.

Results:

Part of the process of gaining insight into the results of this study involved performing a quantitative univariate exploratory data analysis. The EDA analysis of this study is that the distribution of the human deaths caused by the big 4 is right skewed with an outlier, centered on a median of 319, with an IQR from 498.2 to 225 (Figure 1; Table 1). I chose to use the median and IQR as measures of center and dispersion for this EDA because of the presence of an outlier and the skewness of the distribution.

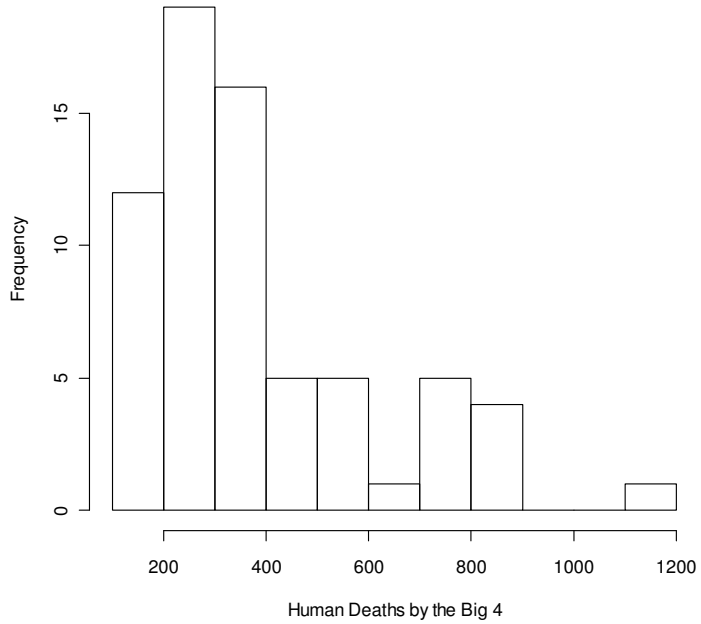

Figure 1: Histogram of Human Deaths from the Big 4 for the years 1940 – 2007.

Table 1: Summary of the statistics of Human Deaths from the Big 4 for the years 1940 - 2007

n	NAs	Valid n	Mean	St. Dev.
68.00	0.00	68.00	384.41	221.05
Min.	1st Qu.	Median	3rd Qu.	Max.
130.00	225.00	319.00	498.20	1135.00

The second step of analyzing the study data was the completion of an 11-step hypothesis test.

1.) A quantitative variable was measured on variables from one population; σ is unknown; thus a 1-sample t-test is required.

2.) The null hypothesis will be about μ and it will be tested against a specific value, namely μ=300. Thus H_o: μ=0 and H_a: μ>0; if H_a is supported then greater than 300 people are killed each year by the big 4.

3.) As I stated**,** α should be set at 0.01.

4.) The data appeared to be part of an observational study; the meteorologists who collected this data did not impart any conditions on whether or not the people died because of the severe weather events. There is no evidence that randomization was used.

5.) The sample size is greater than 40. Therefore, the test statistic given below, in step 6, should reasonably follow a standard normal distribution. In addition σ is unknown.

6.) The hypothesis is about μ. Therefore we want to calculate $\bar{x}$**,** which = 384.41 deaths (See **Table 2**).

7.) The t test statistic is **t**=3.149. (See **Table 2**)

8.) The p-value for this statistic is 0.001224. (See **Table2**)

9.) Reject H_o because p-value < α = .01.

10.) It appears that there is evidence that more than 300 people are killed each year by the big 4.

11.) One is 99% confident that the number of people killed each year by the big 4 is on average at least 320.524 (See **Table 2).**

Table 2: 1-sample t-test results for how many deaths were caused by the big 4 from 1940-2007.

```
        One Sample t-test
data:  w2
t = 3.1489, df = 67, p-value = 0.001224
alternative hypothesis: true mean is greater than 300
99 percent confidence interval:
 320.5235      Inf
sample estimates:
mean of x
 384.4118
```

Discussion:

To reiterate, from step 10 of the 1-sample t-test that there is evidence that more than 300 are killed each year by the big 4. This data appears to support my pre-research premonition that there are more than 300 people killed each year by the big 4.

This information could be used in a variety of ways. First off it could be used as evidence to the government that we need to continue spending all of this money for research on these four main severe weather events. Researchers could show this to the government and say that since there are so many people dying each year due to these events**,** we need to spend more money on finding out ways to better predict where and when these events will happen and also how to better design homes and buildings to withstand these events (mainly hurricanes and tornadoes). This information could also be used to determine where not to build a home or business in the case of a flooding. It could also be used as an eye-opener to the public. The average person probably doesn't know how many people actually dies each year due to these events; they only care if they affect them or someone they know. But, if they could see this maybe, it will open there eyes a little bit to see what important research people are doing to better warn them about oncoming events and how to better build their homes and protect themselves during these events.

To improve upon this study, future research should take place over a longer period of time in order to see if the number of deaths starts to decline over the years. As better designed homes are built and we come up with better technologies to help predict where and when these events will occur**,** you should see the number of deaths decline. Thus proving that all of the research money is paying off and lives are being saved. To further improve, you could also compare where the most weather related deaths are occurring and figure out where more of the money should be spent on research and development. A third and final adjustment that you could make to improve this study is to take out some of the years that don't quite fit in, like the outlier in the year in which hurricane Katrina hit and caused over a thousand deaths in 2005, causing an outlier and swaying our results a little higher. Little things like that could make a big difference and should be considered if this research is done again in the future, when there is more data available.

Sources:

http://www.weather.gov/os/hazstats/images/67-years.pdf

http://www.weather.gov/os/hazstats/sum07.pdf