MAT 1372: Statistics with Probability

Yingci Wu

First Draft

The topic that I became most interested after reading both David Brook articles is that big data might completely fails to predict the weather accurately. This topic is relating to Brooks claims directly, and according to the result of my study about data, I found that data can be very helpful if people know how to analyze it carefully and apply it in appropriate situations. Data can be useful in our daily lives because we human are obsessed with quantifying everything into measureable units. However, sometimes we might not be able to get the desired results even if we have enormous amount of data. In a David Brooks’ column, “The Philosophy of Data,” he claims that data performs well in two areas: First, data “is really good at exposing when out intuitive view of reality is wrong” and “Second, it can illuminate patterns of behaviors we have not yet noticed” (Brooks, 2013). Yes, it is beyond doubt that data can be extremely useful if one uses it correctly; either the data is cross-sectional or just a simple point, it can lead many possible paths which would assist the researcher to assess his hypothesis. However, in his other column “What Data Can’t Do,” Brooks also suggests that “big data does poorly” in many aspects that it cannot always handle big problems, and it sometimes might obscure the reality (Brooks, 2013). I will discuss both articles and my study more deeply in below, also expanding more information based on this topic.

We are living in an era which we depend much of our daily life on information. For example, weather prediction is one of the most important issues in the contemporary society and we all rely to some extent on accurate weather forecast. The methods that scientists use to predict the weather employ a massive amount of empirical data collected globally twenty-four hours around the clock. After the meteorologists gather the statistical data, they would plot the data on a map, run the computer models and finally release the results to the public. Although the process is very objective and scientific, the results may not be always be correct; sometimes the model may not accurately foretell the subsequent weather because the data may be erroneous or the model contains too many unknowns. In the article “Why can’t scientists accurately predict the weather,” Harris is dismay that “How can we put a person on the moon or foretell planetary alignments years in advance, yet still fail to put together accurate weather forecasts?” (2012). Meteorologists understand very well why they are unable to predict the weather accurately even with the aid of supercomputers; they feed into the supercomputers models that evaluate several different large-scale phenomena, each governed by multiple variables, and a small change in any of these variables would profoundly affect the future weather. This also could make the weather condition change over time and thus give a wrong prediction even with all the available data (Harris, 2012). He also claims that this happens because of chaos theory, which states that in a set of highly complex system, small changes in initial condition can produce a large unexpected change in the final results (Harris, 2012).

Even though we all believe on the fact that data can help and clarify the superstitious; however, our attitude toward data can limit our perceptions and thoughts. For example, we may blame forecasters for their inaccurate weather forecast even though we are well aware that it is hard to predict weather even with a supercomputer from National Centers for Environmental Prediction. According to the article, “The Weatherman is Not a Moron”, by Nate Silver in 2013, he argues that after expert meteorologists receive weather forecasts data generated by supercomputers, they would interpret, analyze and compare them with other models with their colleagues in order to make their prediction as accurate as possible. Silver also argues that weather has two additional properties; firstly it is “nonlinear, meaning that it abides by exponential rather than by arithmetic relationships” and secondly, it is also “dynamic” which would make forecasting even more difficult by just using data. Thus the behavior at one point in a certain time can affect its behavior at other time (Silver, 2013). For example: a study in Kansas City found that when a forecaster said there was 100 percent chance of rain over a day, in fact, it failed to rain at “all one-third of the time” (Silver, 2013). People might not care when a weatherman predicts rain and it turns out to be a sunny day. But if it rains when it isn’t supposed to be, guess what? They may accuse the weatherman for providing wrongful information.

Overall, there is not one best way to predict weather accurately because there are still many unknowns the meteorologists don’t know about the weather system, and even if they have all the available data, their effort may still be furtive. Nevertheless data can give us “wonderful ways to understand the past and present” (Brooks 2013). Data is not the answer; they are only information, facts which provide aids to people who need them. Many experts recognize that data on its own isn’t good enough because they have to know when the data can help. And as my research reveals when we are predicting something as dynamic as weather, we really need to decide what data should be collected, and we then could decide when we are supposed to provide it and how it works.

Works Sited

Harris, William (2010, July 07). “Why can’t scientists accurately predict weather?” *Science.Howstuffworks.* Retrieved March 29, 2013. Web:

http://science.howstuffworks.com/nature/climate-weather/atmospheric/scientists-predict-weather.htm

Silver, Nate (2012, September 7). The Weatherman Is Not A Moran. *New York Times.* Retrieved March 30, 2013. Web: http://www.nytimes.com/2012/09/09/magazine/the-weatherman-is-not-a-moron.html?pagewanted=all&\_r=0