Note on Bodacious Biostatistics

C'Asia Griffin, Sujin Kim and Mulatu Lemma

Department of
Mathematics
Savannah State
University
USA

Abstract

Biostatistics is the branch of statistics that deals with data relating to living organisms. Biostatisticians help answer pressing research questions in medicine, biology and public health, such as whether a new drug works, what causes cancer and other diseases, and how long a person with a certain illness is likely to survive. Biostatistics is the connector between math and science as it is a field that merges passion and skill with biomedical science and mathematics and statistics and central to all of science, because science needs that gathering of evidence and the evaluation of that evidence to make a judgment. Biostatisticians use their quantitative skills to team up with experts in other fields, from biologists and cancer specialists to surgeons and geneticists. Most importantly, biostatisticians are not just individuals who deal with numbers only. They play pivotal roles in designing studies to ensure enough data and the right kind of information are collected. Then they analyze, evaluate and interpret the results – accounting for variables, biases and missing data along the way.

Introduction

The role of biostatisticians are to identify and develop treatments for disease and estimate their effects. Once they identify risk factors for diseases, they are able to interpret, analyze, design, monitor, and report results of clinical studies. They also develop methods to answer questions arising from medical and public health data. Lastly, statisticians locate, define and measure the extent of a disease. Ultimately, these objectives help improve the health of individuals and

communities.

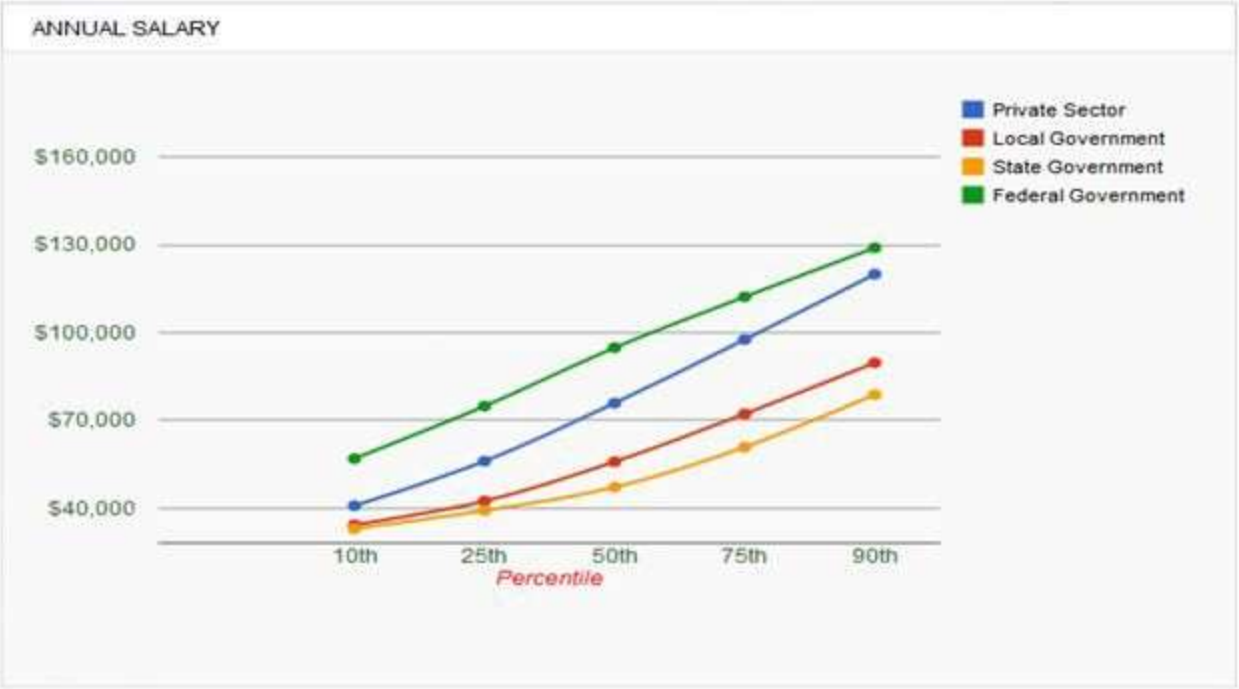
To become a biostatistician, you must have at least a bachelor's degree in statistics, biostatistics, or possibly mathematics. Most jobs in the field of public health will require one to have a master's degree. A master's degree in public health (MPH) with a focus on biostatistics will put one on a solid platform in this career field. About 48% of employed biostatisticians have a Master's Degree.

Duscussion

Biostatistics is a great field if one is interested in impacting the health field or public health without directly working with patients or having to treat patients. A career in biostatistics will allow one to be an important part of the health care and public health systems, and to have a strong impact on the health of many communities in the country and world. This career field also is attractive due to its variety, as one will have many different types of studies to work on.

Many biostatisticians work in public health organizations, government agencies, research companies, and sometimes medical device manufacturers and pharmaceutical companies. It takes many skills and qualifications to be a biostatistician. These skills include strong mathematical skills, statistics skills, problem solving skills, adaptability, written and oral communications skills, strong teamwork skills, critical thinking, and strong computer background.

Most biostatisticians earn a good salary particularly if one holds a master's degree in statistics or an MPH with a concentration in biostatistics. Indeed.com states that the median salary for biostatisticians is $90,000. The Bureau of Labor Statistics states that the median salary for statisticians is $72,000.



According to the Bureau of Labor Statistics, job growth for statisticians will be about 14% by 2020, which is slightly faster than average when compared to other fields. However, economists believe that growth in the biostatistics and health field will be faster because of greater demand for healthcare services. The Bureau of Labor Statistics, also states that there are 25,000 statisticians employed in the United States as of 2012, and there will be about 3,500 new jobs in the field by 2020. The field of biostatistics is emerging with constant growth and opportunities.

I was first introduced to Biostatistics and what it was through my research mentor, Dr. Sujin Kim. I was walking through Herty Hall one day and noticed a flyer stating that the National Institutes of Health (NIH) was recruiting mathematics major who were interested in pursuing a PhD in biomedical sciences. I quickly decided to look more into the opportunity and eventually applied for the scholarship. I had multiple sit down sessions with Dr. Kim where she

basically explained what the field of Biostatistics was, how it was used, how she enjoyed earning her PhD, and the many ways her degree could be used. Through these multiple sit downs with her, I was able to learn that she was a graduate of The University of Iowa. This awakened my view on institutions farther away from home and how they may be beneficial to me.

After completing an interview process, I was granted the MARC U-Star Scholarship from NIH. One of the major requirements of the scholarship is to complete research during the summer months, the first year on campus and the second year off campus at another college. My first summer, I began researching cervical cancer and its correlation to marital status. During this time, I was able to learn about chi-square, odds of ratio, p value, and many more other concepts.

I realized at this point how challenging the subject matters can be, even if in actuality they are very simple. Conducting research for the first time was a rewarding experience where I was able to acquire many pros and cons to use towards future research endeavors. My complete research project was fascinating to say the least. I was able to support my hypothesis with data evidence to show that individuals who were not married were more likely to contract cervical cancer than individuals who were married.

Below is an excerpt from my first ever written research paper:

"Although ranked 14th in frequency among all cancers in United States women, cervical cancer is the most common cancer among reproductive organ cancers in women. It holds an incidence rate of 8.1 cases per 100,000 women per year in the United States with a mortality rate of 2.4 deaths per 100,000 women per year. Mostly all cases of cervical cancer are caused by a type of human papillomavirus (HPV). NIH is currently supporting efforts to make cervical cancer prevention, screening, and treatment more affordable to help reduce cervical

cancer incidence and mortality in developing nations and to make interventions more cost-effective in the United States.

According to the Centers for Disease Control and Prevention, HPV is the most common sexually transmitted infection in the United States. Seventy-nine million Americans are infected with HPV. At times, it is difficult to determine when an individual is infected with HPV because it can take years for symptoms to begin developing and showing. However, not all cases of HPV lead to cancer as some situations just go away over time on its own. This is why HPV testing intervals are not recommended to be close together, such as annually, because the findings of an abnormality may not necessarily be a long term matter that will lead to cancer or other diseases, but more so a temporary change.

Pap testing is important in the fact that it detects abnormal cells which may later develop into cancer. For unfortunate individuals who may already have cancer and not be aware of it, pap testing help to identify these cancer cells. Through recent years, the liquid-based pap testing has increased in popularity and productivity over the conventional method of pap tests. This method allows the same sample tested for abnormal cells to be used for the testing of high risk types of HPV.

The importance of women actually going to get HPV and pap testings is stressed widely. This is recommended for women aged 21 and older. After all, this will allow more and more individuals to detect cervical cancer and other diseases at an earlier stage. If the situation is fortunate enough, the abnormalities can be detected and treated before cancer even develops.

The relationship between Pap and/or HPV test results and legal marital status depends largely on an individual's current relationship status. HPV is estimated to be the most common

sexually transmitted infection in the United States. Previous papers have found a correlation between pap and HPV test results and the amount of partners an individual has. These data sets have been separated to categorize the partners into a marital status. There is expected to be a positive correlation between more partners and having cervical cancer.

In a previous study, it was found that 93% of cervical cancers worldwide contain HPV. This finding attributes to be "the highest worldwide attributable fraction so far reported for a specific cause of any major human cancer." Being that this number is considered to be underestimated, there were extremely rare cases where HPV was not found. From this study, it is clear that HPV plays a major role in the development of cervical cancer. In our study, we will further investigate the direct effect HPV testing has on cervical cancer.

Categories of marital status include never married, former married, and current married. The subsections of former married has factors of the individual being divorced or widowed. These categories were compared to Pap and/or HPV test results of both normal and abnormal. With these categories, we assume that the risk of cervical cancer is constant, no matter the variable. If variables used to define the categories are inaccurate, it will cause deficiencies within our data set. With individuals who "Don't Know" or "Refused" their marital status included in the data set, the relative risk associated with cervical cancer is difficult to determine because it is unclear whether or not the individual has been with a partner. Other continuous variables that may be important include an individual's weight, frequency of fruit and vegetable consumption, frequency of beans and whole grain consumption, vitamin usage, cigarette use daily, and amount of live births.

In addition to these issues, some studies incorporate data from all individuals including males, failing to realize cervical cancer is a type of cancer that is only found in women. However, when genetic counseling plays a role in the study, such as genetics in immediate family members, it may be important to compare how different marital status perform within the study population. Our comparison includes 7,709 participants with abnormal and normal testings of pap and/or HPV. In marital statuses, we use never married, former married, and current married. We compare the performance of three separate categories using data from an immense cancer sample in adults. An appropriate measurement involving less partners and other positive factors can play a significant role in the prevention of contracting many diseases and illnesses, including cervical cancer.

Using SPSS, we found the cumulative frequency of every category before and after removing the extraneous categories (known as data cleaning). With the cross tabulation, one is able to see the entire group of individuals represented and the data gathered from them. Using a 95% confidence interval, several types of odds of ratios were calculated to determine the likelihood that one will develop cervical cancer based on the relationship between marital status and pap/HPV testing. We also used binary logistic regression to see how marital status would determine an outcome. Lastly, we found significance levels (p-values) to accept or reject the null hypothesis.

In the PapHPV results, the frequency for abnormal testings was 1202 with a 15.6 among percent, valid percent and cumulative percent. The frequency for normal testings was 6507 with 84.4 percent and valid percent, and a 100.0 cumulative percent.

In MaritalStatus results, the frequency for never married was 1691 with a 21.9 among percent, valid percent, and cumulative percent. The frequency for current married was 6018 with a 78.1 among percent and valid percent and 100.0 cumulative percent.

In the MaritalStatus * PapHPV Crosstabulation, 305 participants were never married and had abnormal testings. 1,386 participants were never married with normal testings. 897 participants were currently married with abnormal test results. Lastly, 5,121 participants were currently married with normal test results.

In the tests of conditional independence, Chi-Squared Cochran's was 9.835 and Chi-Squared Mantel-Haenszel was 9.597.

The Mantel-Haenszel Common Odds Ratio Estimate was 1.256.

In the Model Summary, the -2 Log likelihood was 6664.041 with a .001 Cox & Snell R Square and a .002 Nagelkerke R Square.

We formed the equation of the PapHPV Test Results from the "Variables in the Equation" table with the slope of the line at -.228 and the constant exponent being 1.742.

PapHPV Test Results = -.228 * MaritalStatus + 1.742

MaritalStatus is significant with p-value of .002.

Odds of ratio is .796.

Never married participants have a lower odds of having a positive result on PapTest/HPV Test than current married participants.

The -2 Log likelihood has a great reliability of producing the outcome expected.

Our participants were adults aged 18 and over. We do not know the exact age of participants nor the number of partners they may have. For future studies, we would like to look

further into individuals living in areas where having more than one partner is normal. (Will the chances of having cervical cancer increase?)"

I was able to expand my knowledge on the topic by investigating future suggested studies for the semester of Fall 2018. I explored factors affecting cervical cancer which were individuals who consumed red meat, processed meat, and smoked cigarettes. This portion of the research was very alarming because I consume red meat and processed meat daily. To know that something as small as consuming meat can affect my overall health for a lifetime is scary. Based on our results, we did not find a big enough difference between Model 1 and Model 2 to report and consider the data statistically.

Upon completing my research, knowing Dr. Kim had attended, went through, and successfully completed the process at the University of Iowa made me actually consider attending there. With much thought, I eventually applied to the summer program at the University of Iowa. With the help of a great resource, connection, and recommendation, I was accepted into the Iowa Summer Institute in Biostatistics program at the University of Iowa for the summer of 2019.

During the seven week program, I was able to realize that statistics extends far beyond the classroom. Through my topic of how heat waves affect criminal activity and its association, I was able to see that statistics is not just mathematics. This weather and climate issue among criminal activity intrigued me to the realization that a large aspect like weather is affected by daily calculations. Although this topic seemed as if it would not correlate to numbers at all, there were actually several times where we had to perform complex and complicated calculations to achieve the results we hoped to see.

Our goal was to assess the association between heat waves and criminal activity. My group partners and I originally believed that criminal activity would increase as more heat waves were present. Using our former knowledge, we knew that heat triggered the aggressors in our bodies. So we assumed that as individuals became more aggressive during heat waves, the more crimes would occur. For our ending result, we actually found the complete opposite. We found that on days where heat waves were not present were the days where criminal activity increased the most. After analyzing our results, we contributed this happening to the fact that as it gets hotter outside, the less people actually want to be outside. Individuals are more likely to stay inside during periods of heat waves altogether.

There are not many biostatistics graduate programs offered in the United States of America. I have taken interest particularly in The University of Iowa, Emory University: School of Medicine, University of Alabama at Birmingham, and the University of Georgia. I know that regardless of my decision, the program I choose will aid me in becoming the successful biostatician I aim to be. These programs will require hard work, but I know that I can do it. Having a first hand look and experience at the University of Iowa's campus, department of biostatistics, and overall environment allowed me to grow to love the students, faculty, and surrounding areas.

When chosen for a graduate program, I will cherish the knowledge I have acquired from being in the summer program and how to properly pace myself to get through the complex topics. I am also aware that depending on the program and the university, the workload will be very strenuous. Some departments are all about numbers and how that makes their higher institution look. Schools like those are to be highly avoided because they do not really care about

their students. In reality, they want their students to be in competition with one another. It is basically the best example of survival of the fittest.

When students are put against one another, it never ends well. I want to find a program that will enjoy and appreciate my efforts for itself. I do not want to cause unnecessary stress on myself just because I am worried about being a part of the top 1%. I truly do believe that pressure does indeed creates diamonds, but sometimes there are more efficient ways to accomplish success.

Biostatistics extends beyond just the field into other fields of medicine, health, nutrition, epidemiology, genetics, biology, and even art. I find this to be so fascinating because this field is not widely referenced or known about. To know that it correlates to so much more than just itself is nothing short of amazing. This is why I chose to pursue a higher degree in Biostatistics. I like the fact that I am able to be unique and different in not only the subject I chose to study, but the material that I will be learning. There are not many biostatisticians in the world, so to know that I will be among the few amazes me and pushes me even more to never give up when school gets hard.

Pursuing a higher education is very important to me. My current academic success has only encouraged me so much more to complete my undergraduate degree and attend graduate school. In my opinion, education is the key to success and the gateway to endless opportunities. I am what they deem as a first generation college student, so I definitely want to break down barriers and statistics to achieve two of the most important things to me. I come from humble beginnings which taught me to persevere through any obstacle and struggle I ever endure.

The process of obtaining my PhD in Biostatistics will require another five years in school. Sometimes this can be very discouraging with the way I already feel about school. I am completely over the process of waking up early mornings, giving my all just to earn terrible grades in return. Hopefully, changing my familiar surroundings and replacing them with new territories will open up my mindset and refresh my motivation and spirit towards earning a second degree. No matter the circumstances, I know that I will one day be an achieved biostatistician by the name of Dr. C'Asia De'Nay Griffin.

References

Biostatistician Public Health Career, Salary & Job Description. (n.d.). Retrieved December 9, 2019, from https://www.onlinemphdegree.net/biostatistician/.

Department of Biostatistics. (n.d.). Retrieved December 8, 2019, from https://www.biostat.washington.edu/about/biostatististics.

Jack, J. (2014, December 30). Application of Biostatistics. Retrieved December 9, 2019, from https://www.slideshare.net/jippyjack5/application-of-biostatistics.