Exploring Epidemiology on Campus

Sara Grassel, Jessica Kraker (krakerjj@uwec.edu)
Mathematics Department, University of Wisconsin – Eau Claire
Student/Faculty Research

Introduction

Epidemiology is "the study of the distribution and determinants of disease in human populations" [1]. The study of Sexually Transmitted Infections (STIs) and their modes of transmission is an important segment of epidemiology. In terms of education and information for the student body, the study and assessment of the sexual health of a campus population is a step towards the goal of improving the quality of the students' college experience. This project provided a service of statistical analysis to UW-Eau Claire Student Health Service in answer to that purpose.

Collaboration with Student Health Services

Assessment was performed by Student Health Services (SHS) during April 2007; a survey designed by the American College Health Association (ACHA) was administered to a random sample of currently enrolled students. Out of this sample, there were 337 students who responded. Throughout the duration of our research, Laura Chellman, Director of SHS, provided us with the necessary information to allow appropriate analysis to be completed. Of particular interest, we wanted to examine behaviors associated with sexual health and how to advocate such behaviors. This is a stated goal of UW-Eau Claire's SHS.

Statistical Process

All steps in the diagram below are vital to the process; research and practical work were involved in each step.

Formulate Research Questions → Design & Administer Survey* → Compile & Validate Data → Choose Appropriate Method to Analyze → Inference

*This step was performed by ACHA and Student Health Services.

Example Process

Research Question: Is there an association between whether or not someone was ever diagnosed with depression (#41) and whether or not stress affected his/her performance in the last school year (#44x)?

The figures to the right and above show the actual survey questions [2] asked to each participant; the responses were then used to address the research question. The data was summarized, checked for missing information or errors, and then tabulated (see **Table 1**).

The variables from the previous research question are both categorical. From **Table 2** one can see an appropriate test for this data is the Chi-Square Test for Association. The test was run in the computer program *Minitab*, and the summary is provided below.

Table 1

	Diagnosed with Depression			
Stress		Yes	No	Total
Affected	No	21	128	149
Academic	Yes	38	91	129
Performance	Total	65	270	335

Test summary

H0 : Row and Column variables are independent
HA: Row and Column are not independent

$$\chi^2 = \frac{(21-31.62)^2}{31.62} + \frac{(128-117.38)^2}{117.38} + \frac{(38-27.28)^2}{27.28} + \frac{(91-101.62)^2}{101.62} = 9.761$$

DF = 1, Chi-Sq = 9.761, P-Value = 0.002

We conclude that there is sufficient evidence [P-Value < 0.01] to suggest that there is an association between whether someone was ever diagnosed with depression and whether or not stress affected his/her academic performance in the last year.

References

[1] "A Study Guide to Epidemiology and Biostatistics", Morton, R. F., Hebel, J. R., and McCarter, R. J.; Aspen Publishers, **2001**.
[2] http://www.achancha.org/docs/SampleSurvey_ACHA-NCHA_Sp00-Sp08.pdf
[3] "Handbook of Parametric and Nonparametric Statistical Procedures", Sheskin, D. J.; Chapman & Hall / CRC, **2000**.

Data Comparison Methods

Once the data to address the various research questions were summarized, we determined what tests were appropriate based on the type of variables used. The variables were numerical, ordinal (values had inherent order), or nominal (values had no inherent order). When comparing different types of variables, only certain tests may apply; see **Table 2** for some relevant methods (see also , for example, reference [3] for descriptions and applications of these methods).

Table 2

Methods Used	Variable Types
Gamma Kendall's Tau-C Spearman's Rho	Two Ordinals for Association
Chi-Square	Two Categoricals for Association
Logistic Regression	Predictor Numeric & Response Categorical
Ordered Logistic Regression	Predictor Numeric & Response Ordinal
Somer's D Gamma Kendall's Tau-C Spearman's Rho	Predictor Ordinal & Response Ordinal
Lambda Cramer's V	Predictor Categorical & Response Categorical
T-test Mann-Whitney	Predictor Categorical & Response Numeric

Selected Results and Conclusions

UW-Eau Claire Health Services can use the analysis and summary provided by our work in order to pinpoint areas of improvement in the sexual health of students. By raising awareness of these issues, the student body can positively react and hinder the issue from reoccurring. Overall, this research allowed Sara to not only use her math experience, but also provided a service to the UW-Eau Claire campus. Below are listed a few of our research questions (with references to the survey questions) and conclusions, with inferences applicable to the appropriate population (UW-Eau Claire students who would respond to such a health survey).

Is there an association between the number of alcoholic drinks someone had the last time he/she partied (#13) and whether or not they had unprotected sex due to drinking within the last school year (#18g)?
Based on a logistic regression analysis there is strong evidence [P-Value <0.01] to suggest that the number of drinks someone had the last time he/she partied is linearly associated with the logs odds ratio of whether or not they had unprotected sex due to drinking within the last year, through a linear model. Therefore, the number of drinks someone had the last time he/she partied can be used to meaningfully estimate the proportion of individuals who had unprotected sex due to drinking within the last year.

Is the number of partners an individual has had in the last school year different than the number of partners they perceive the typical student has?
There is sufficient evidence to suggest [P-value <0.01] that the difference between the number of reported partners to the number of partners that the individual perceives the typical student has is not 0. We are 95% confident that the true mean difference between reported and perceived number of partners lies between -1.8423 and -1.4610. In other words, the perceived number of partners a typical student has is greater than the reported number of partners, on average.

Is there an association between whether or not a person received information from his/her college or university (#2g) concerning pregnancy prevention and whether or not they had vaginal intercourse within the last school year (#30)?
There is not enough evidence to suggest that dissemination of information about pregnancy is based on sexual activity. Thus it is reasonable that there is a general dispersal of information to students regarding pregnancy.

Is there an association between the number of partners someone had sex with within the last school year (#20) and the location in which they live (#54)?
Based on a nonparametric test for median difference, there is significant evidence [P-value<0.001] to suggest that median number of partners differs by whether or not the student lives on campus. Visual tools show that the distribution of number of partners is higher for students living off-campus.

Acknowledgements