Gender Demographics in Introductory Statistics and Other Sciences at UTEP

Rocio Estrada
restrada18@miners.utep.edu

Gabriel Pardo
real_grpardo2@hotmail.com

The University of Texas at El Paso

STAT 2480 Final Project

Andrew Dynneson
dynnesonTeaching@gmail.com
Introduction

In the math and science fields of study there has always been a larger male to female ratio\(^5\). Studies have shown that there has been an increase of women continuing on into higher education in the math and science fields. In 1993, 23.8% of mathematics and statistics doctorate degrees were earned by women, while in 2012, 28.2% of doctoral graduates were women\(^6\). Women account for 38.3% of those 72,000 statisticians, according to the most recent data of the US Department of Labor\(^7\). With the increase of women holding positions as statisticians there is some belief that there will be more women entering into the STEM fields\(^5\). Women also accounted for 44% of Ph.D. graduates in statistics/biostatistics in 2013, compared with the 27% of doctoral graduates for other mathematical sciences\(^7\). Furthermore, it has been observed that women are actually a majority in biological sciences, more than 70% of graduates of every educational level are women\(^7\).

In the University of Texas at El Paso, core curriculum courses are required for every major, and these courses vary depending on the degree plan. The data to be studied is the ratios of females to males taking statistics courses at UTEP. It has been observed that more females are enrolled in Elementary Statistics 2480 in Fall 2015. The Chi-Square Test was implemented in this study to test the Homogeneity. It was also observed that STAT 2480 is, in some degree plans, under the section of core curriculum courses\(^9\). Observed numbers for STAT 2480 has a higher female ratio, which coincides with the female enrollment at UTEP. Although, when students declare their major and select other upper level elective courses like STAT 3320, the ratio changed. Null and alternative hypotheses are set up for this study as follows:

Ho: There is no difference between genders in enrollment of a particular course when compared to the student population at UTEP.
Ha: There is a difference between genders in enrollment of a particular course when compared to the student population at UTEP

Through the use of the Chi-Square Test, a p-value will allow for either rejection of the null in favor of the alternative hypothesis, or non-rejection of the null. Additionally, the p-value determines which gender has a higher number enrolled per class. With the data gathered, we conclude that females have a higher number of individuals enrolled in the STAT 2480 course, when compared with the UTEP student population. This could be due to core curriculum requirements; for example, the requirements of the Biology, Forensic Science, Mathematics (with Statistics minor), Environmental Science (with Biological or Geological concentration) or Cell and Molecular Biochemistry degree plan would require. It is not the same case for CHEM 1305 since there are more males than females enrolled. We have also included the elective Probability and Statistics 3320 in order to compare ratios of enrollment.

From information obtained from the United States Census Bureau, we know that there are more females than males in both El Paso County and Texas. In El Paso County, males occupy the 49.0% of the population, while in Texas 49.6%. A higher female ratio is also true for general UTEP population, which since early 2000s shows a majority of female student ratio when compared to male ratio. Throughout 2004-2007, there have been 55% female and 44% male enrollment, fluctuating only by decimals. Furthermore, in 2014 female enrollment was 7.96% more than male enrollment. Prior, in 2013 the same pattern continues with 53.98% of females and 46.02% of males in UTEP enrollment, all these including all undergraduates and graduates.

The purpose of this project is to examine the male and female ratios in statistical courses and compare the information with the current 2015 demographics of other courses and general
UTEP undergraduate population. We expect to find results that associate Elementary Statistics enrollment with either continued statistical education or core curriculum requirements.

The variables associated with the research include quantitative values of number of people enrolled in courses and qualitative variables by categorizing into males or females. Data of study is observational. Demographic data was obtained from the Administrative office CIERP that agreed to provide our group with this information. The sample of study is chosen to be Elementary Statistics 2480. The female population of UTEP, El Paso and Texas is already known to be higher than the male population. A lurking variable to be considered includes majors other than mathematics/statistics that require STAT 2480 as core curriculum, for example biology or forensic science. This could be an important variable since elementary statistics is a requirement for such degree plans meaning that female students are considering a different carrier path other than statistics.

Data

The data for this study was obtained through several sources. UTEP CIERP department provided the information needed for the amount of students enrolled for STAT 2480 and 3320, BIOL 1305, CHEM 1305, and general UTEP population. Other statistics were obtained in different official Internet websites. Table 1 represents the populations provided by CIERP. Included in the table is the comparison between male and female population in El Paso and Texas.

Table 1 Male and female populations
### Fall 2015 Population

<table>
<thead>
<tr>
<th></th>
<th>Females</th>
<th>Males</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stat 2480</td>
<td>166.00</td>
<td>86.00</td>
<td>252.00</td>
</tr>
<tr>
<td>Stat 3320</td>
<td>36.00</td>
<td>40.00</td>
<td>76.00</td>
</tr>
<tr>
<td>BIOL 1305</td>
<td>806.00</td>
<td>435.00</td>
<td>1,241.00</td>
</tr>
<tr>
<td>CHEM 1305</td>
<td>356.00</td>
<td>358.00</td>
<td>714.00</td>
</tr>
<tr>
<td>UTEP UG</td>
<td>10,862.00</td>
<td>9,358.00</td>
<td>20,220.00</td>
</tr>
<tr>
<td>El Paso</td>
<td>425,078.00</td>
<td>408,409.00</td>
<td>833,487.00</td>
</tr>
<tr>
<td>Texas</td>
<td>13,586,307.00</td>
<td>13,370,651.00</td>
<td>26,956,958.00</td>
</tr>
</tbody>
</table>

### Statistical Methods

Chi-Squared was used for hypothesis testing and Excel software was used for calculations. The proportion of males and females at UTEP was obtained by dividing gender by the total UTEP population. Expected values of each course were calculated by multiplying the proportion of gender at UTEP by the observed gender ratio. The chi-squared formula was then used:

\[ X^2 = \frac{\sum \text{(observed value-expected value)}^2}{\text{expected value}} \]

Since 2 sample variables are used, 1 degree of freedom is used to obtain p-values from the \( X^2 \) distribution table.

### Results

In Table 1, it was observed that females had a higher number of individuals enrolled in the undergraduate classification at UTEP. However, when observed by course, the population of females in the BIOL 1305 course was substantially larger than the population of males. Similarly, In STAT 2480 the female population was nearly twice as much as the male population, which was a remarkable with a p-value that was less than the 5% significance level.
Contrastingly, in CHEM 1305 there were significantly more males when compared to the UTEP population. In STAT 3320, there was no meaningful difference in gender when compared to the UTEP population.

The Chi-Square Test was utilized to interpret data. This test was calculated for four different courses’ population of females and males in comparison to the Undergraduate population at UTEP. STAT 2480 and BIOL 1305 had a calculated p-value <0.0005, which results in enough evidence to reject $H_0$. Likewise, CHEM 1305 rejects $H_0$ with a p-value between 0.05 and 0.025. STAT 3320, however, has a p-value of >0.25, so we do not reject $H_0$.

**Discussion**

This study allowed for a close look into the ratio between females and males in both STAT courses and other science courses. As observed in STAT 2480, BIOL 1305 and CHEM 1305, p-value was less than the 5% significance level, which resulted in the rejection of $H_0$. The Chi-Square test performed on STAT 3320 resulted in a p-value >.25, therefore we do not reject $H_0$ suggesting that there was no difference between gender in that statistics class when compared to UTEP population. For STAT 3320, considering that is an elective course and that chi-square does not reject $H_0$, we could speculate that males were more interested in enrolling for that class than females in Fall 2015.

The results of this investigation suggest a larger female to male ratio at UTEP. The study involved Elementary Statistics 2480, which is considered core curriculum (required), while STAT 3320 is optional to satisfy the required credit hours towards a mathematics degree with statistics minor. In order to get a better understanding of the ratios between females and males in
statistics, gender demographics of other required statistical courses could be obtained so further investigation can be performed. Furthermore, gender demographics for mathematical courses could be obtained so a comparison between math and statistics demographics can be done. Additionally, to follow up in our previous speculation indicating more male interest in STAT 3320, data from other schools could be retrieved in order to compare core curriculum and elective statistical course enrollment. New data could indicate which gender is more interested in statistics. Since UTEP does not have a Statistics major, gender demographic data from other schools pertaining to that specific major could also serve for comparison.
Works Cited

   http://research.utep.edu/Default.aspx?tabid=44271

   http://universitycommunications.utep.edu/facts/index.html

   http://admin.utep.edu/default.aspx?tabid=50864

   http://quickfacts.census.gov/qfd/states/48/48141.html

   Labor USA, 04 Feb 2015. Web 02 Dec 2015. https://blog.dol.gov/2015/02/04/women-in- 
   statistics-beyond-the-headline/


Appendices

Table 2

<table>
<thead>
<tr>
<th>Fall 2015 Population</th>
<th>Females</th>
<th>Males</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stat 2480</td>
<td>166.00</td>
<td>86.00</td>
<td>252.00</td>
</tr>
<tr>
<td>UTEP UG</td>
<td>10,862.00</td>
<td>9,358.00</td>
<td>20,220.00</td>
</tr>
<tr>
<td>%</td>
<td>0.54</td>
<td>0.46</td>
<td></td>
</tr>
</tbody>
</table>

Expected Females: 135.3721068
Expected Males: 116.6278932

Chi-Squared: 14.97280504
df: 1
p-value: <0.0005

Table 3

<table>
<thead>
<tr>
<th>Fall 2015 Population</th>
<th>Females</th>
<th>Males</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stat 3320</td>
<td>36.00</td>
<td>40.00</td>
<td>76.00</td>
</tr>
<tr>
<td>UTEP UG</td>
<td>10,862.00</td>
<td>9,358.00</td>
<td>20,220.00</td>
</tr>
<tr>
<td>%</td>
<td>0.54</td>
<td>0.46</td>
<td></td>
</tr>
</tbody>
</table>

Expected Females: 40.83
Expected Males: 35.17349159

Chi-Squared: 1.2328834
df: 1
p-value: >.25
Table 4

<table>
<thead>
<tr>
<th>Population</th>
<th>Females</th>
<th>Males</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIOL 1305</td>
<td>806.00</td>
<td>435.00</td>
<td>1,241.00</td>
</tr>
<tr>
<td>UTEP UG</td>
<td>10,862.00</td>
<td>9,358.00</td>
<td>20,220.00</td>
</tr>
</tbody>
</table>

% 
- expected females: 0.54
- expected males: 0.46

chi-squared: 29.12655791, 62.93428093
- df: 1
- p-value: <0.0005

Table 5

<table>
<thead>
<tr>
<th>Population</th>
<th>Females</th>
<th>Males</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHEM 1305</td>
<td>356.00</td>
<td>358.00</td>
<td>714.00</td>
</tr>
<tr>
<td>UTEP UG</td>
<td>10,862.00</td>
<td>9,358.00</td>
<td>20,220.00</td>
</tr>
</tbody>
</table>

% 
- expected females: 0.54
- expected males: 0.46

chi-squared: 1.979483975, 4.277106858
- df: 1
- p-value: 0.05>p-value>0.025
Statistics 2480 Research Proposal Check List

_____ Cover Page.

_____ TITLE: Choose a descriptive title for your research project.

_____ INVESTIGATORS: List the names and emails of all the student researchers who will be working on this project. Statement of the Problem: This section should be written in paragraph form using complete sentences.

_____ Describe the proposed research. Include a clear statement of the research question.

_____ Provide background information.

_____ Justify the significance of the research, that is, Why is this worth studying?".

_____ Clearly define the population to which the findings of this research will apply.

_____ Clearly define the primary endpoints, that is, define the main variables to be collected as part of this study. Are the variables quantitative or qualitative?

_____ Is it a designed experiment, observational study, survey design?

_____ Describe the sampling scheme. Describe any randomization planned, such as random selection of subjects or random assignment of treatments.

_____ How will you determine whether the sample is representative of the population? Clearly define the variables that will be measured in this study for the purpose of determining whether the sample is representative of the population.

_____ Define lurking variables that might be confounded with study, or that could explain associations between the primary response and explanatory variables in this study. Which lurking variables will you choose to measure as part of this study? Proposals Requiring Permission for Use of Human Subjects: For each item, provide the information requested below in the text of the proposal.

_____ Explain whether and how women, minorities, and children (anyone under 21) will be included as subjects in this study. Will anyone interested in participating in the study be allowed to participate? If not, explain who is to be excluded from the study, and why they are to be excluded.

_____ Indicate the specific sites involved in the research study, that is, the locations from which research subjects will be recruited or where observational data will be recorded.

_____ Describe the subjects (exclusion/inclusion criteria), and how they will be solicited. Appendices _____ Attach a draft of the questionnaire (if any) to be used in the research project.

_____ For observational studies, attach tables that will be used to record information observed.

_____ Attach a draft of the proposed informed consent form. (Look at the example on Blackboard) 1. Describe the risks involved in participating in the study. 2. Describe the benefits to be gained by participants (if any) and the benefits that may accrue to society in general as a result of the planned work. Discuss the risks in relation to the anticipated benefits to the participants and society. 3. Describe provisions to adequately protect the rights and welfare of prospective research subjects.