NSSE Data: Differences Between Women and Men at Olin

Molly Farison
October 20, 2010

I. Introduction

This report explores the differences between the learning experiences of female and male students at Franklin W. Olin College of Engineering based on responses to survey questions involving coursework and interactions with faculty. Olin College emphasizes a high female:male student ratio, so gender in engineering education is a pertinent subject of research. The National Survey for Student Engagement (NSSE) poses questions about efforts that students put into their studies and other educational activities, as well as the efforts of colleges and universities to provide learning opportunities.

II. Summary of Data

The data used in this study came from NSSE results of Olin College first-year students in the years 2007-2010. Responses came from 123 female students and 134 male students; this is a small sample size in general terms, but it represents approximately half of the first-year students at Olin in the selected four years. Responses for the fifteen survey questions analyzed in this study ranged from 1 to 4, where 1 = “Never,” 2 = “Sometimes,” 3 = “Often,” and 4 = “Very Often.” This 1-4 scale is appropriate for this self-reported data because survey respondents know the difference between “Often” and “Very Often” more than they would know the difference between a 9 and a 10 on a 1-10 scale.

In order to effectively summarize which survey questions were answered significantly differently by female and male students, I compared the difference between standardized means for the 15 NSSE questions explored in this study. Measuring the difference in a standardized mean (mean divided by pooled standard deviation, or μ/σ) provides a more robust measure of comparison than differences in mean. The four questions with the most significant differences in μ/σ are explored in Table I. The series of NSSE questions addressing coursework and faculty interactions begins with this prompt: “In your experience at your institution in the current school year, how often have you done each of the following?” The four prompts summarized in Table I are as follows:

-“Asked questions in class or contributed to class discussions” (abbreviated as “Class participation”)
-“Had serious conversations with students who are very different from you in terms of their religious beliefs, political opinions, or personal values” (abbreviated as “Serious conversations”)
-“Examined the strengths and weaknesses of your own views on a topic or issue” (abbreviated as “Examined views”)
-“Worked harder than you thought you could to meet an instructor’s standards or expectations” (abbreviated as “Worked hard”)

Summary statistics for these five questions are in Table I.

Mean and mode for these four questions were the same for both genders. For the other eleven questions analyzed, the difference in μ/σ ranged from 0.01 (i.e. “Tried to better understand someone else’s views by imagining how an issue looks from his or her perspective”) to 0.07 (i.e. “Discussed ideas from your readings or classes with faculty members outside of class”). These differences in standardized mean are put on the line for the gender that had a higher μ/σ in Table I.

TABLE I
SUMMARY STATISTICS

Question	Gender	μ	σ	Diff in μ/σ
Class	female	3.04	0.84	
participation	male	3.20	0.78	+0.47
Serious	female	3.14	0.82	
conversations	male	3.24	0.78	+0.34
Examined	female	2.80	0.81	
views	male	2.87	0.83	+0.08
Worked	female	3.03	0.89	+0.08
hard	male	2.96	0.87	

III. Questions that Arise from the Data

Looking at the difference between means as well as the difference between μ/σ, it appears that at Olin College women raise their hand in class less frequently than men. The median and mode of the data are the same for both genders, but the difference in μ/σ is large compared to the other 14 questions analyzed. Several other questions investigated in this study also may show a difference between the experiences of female and male students, but support from anecdotal evidence supports the claim that at Olin College women do raise their hand in class less frequently than men, which is why that question is addressed more than the others in this study. Because Olin College prides itself on its high female:male student ratio among engineering colleges, it is surprising how much of a difference in class participation there is.

IV. Probability Mass Functions (PMFs)

A probability mass function (PMF) shows the probability of each possible event in the form of a histogram. In this study, a PMF shows the fraction of respondents that gave each response. In Figure 1, the pattern that men seem to answer questions in class more frequently than women do is quite apparent.

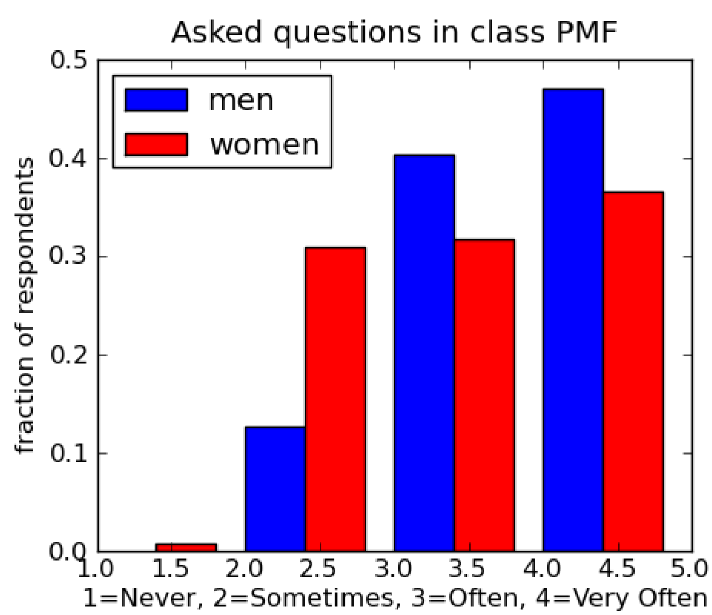

Fig. 1. *This PMF shows that female students report that participate in class less often than male students do.*

In Figure 2, the sum of the women's responses for "Often" and "Very Often" are slightly higher than the sum of men's responses for the same categories. This makes sense given that μ/σ was slightly higher for women than for men (a difference of 0.01).

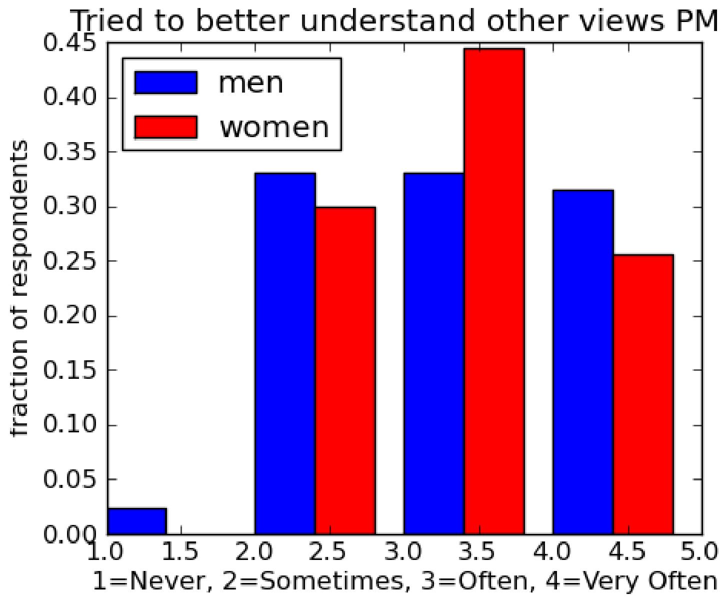

Fig. 2. *This PMF shows that there is not a large difference between female and male responses to the prompt "Try to better understand someone else's views."*

In Figure 3, it looks as though more women than men "worked harder than they thought they could to meet an instructor's standards or expectations," because responses for women are more concentrated in the right half of the graph. This matches up with μ/σ being 0.08 higher for women. A chart like this makes the significance of a difference in μ/σ easy to see.

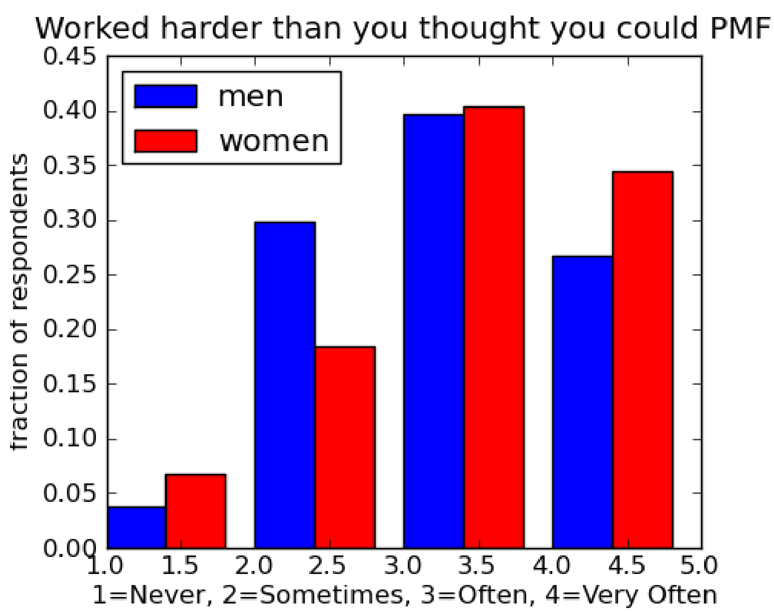

Fig. 3. *This PMF shows that women generally answer higher than men for the prompt "Worked harder than you thought you could to meet an instructor's standards or expectations."*

V. Conditional PMF

In a conditional PMF, a PMF is computed using responses that fit a specified condition. In this study, all of the PMFs shown thus far are conditional in the sense that a response is included in the female PMF only if the respondent is female and included in the male PMF only if the respondent is male. Figure 4 includes only respondents that answered "Very often" to the prompt "Discussed grades or assignments with an instructor." The PMF shows their responses for the prompt "Talked about career plans with a faculty member or advisor." The data suggest that males who discuss classwork with faculty frequently are more likely to take the opportunity to discuss their futures with faculty than women who discuss classwork with faculty just as frequently. However, it is not clear whether this difference in the PMF indicates a pattern in the student body as a whole, especially since this PMF represents the responses of only 32 women and 29 men.

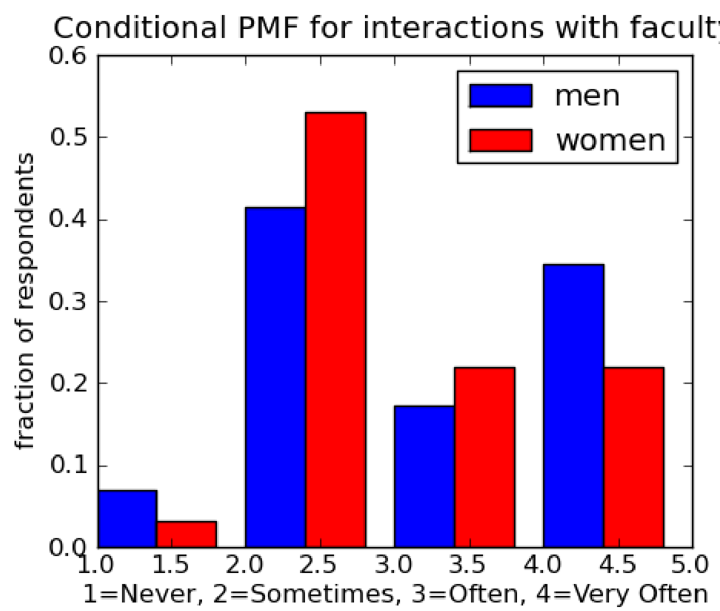

Fig. 4. *This PMF shows results for the survey prompt "Talked about career plans with a faculty member or advisor," with responses taken from respondents who answered "Very often" for the prompt "Discussed grades or assignments with an instructor."*

VI. Cumulative Distribution Functions (CDFs)

A cumulative distribution function (CDF) shows percentile ranks out of 1 as a function of survey responses. In this study, a CDF shows the cumulative fraction of respondents for each possible response.

In Figure 5, it is clear that the vast majority of responses for the prompt "Come to class without completing readings or assignments" range from "Never" to "Often," with very few responses of "Very Often." There were fewer responses of "Very Often" from women than there were from men, which suggests that the most delinquent students at Olin tend to be male. The CDF also shows that approximately the same fraction of respondents of both genders never come to class unprepared. The standardized mean is 0.04 higher for men, which suggests that the difference may be significant, but not as significant as the difference in responses for the prompt "Asked questions in class or contributed to class discussions," which has a difference in μ/σ of 0.08.

VII. Modeling with a Continuous CDF

From the exponential shape of the CDF that is the left-most graph in Figure 6, it looks like results from the prompt "Talked about career plans with a faculty member or advisor" can approximated by

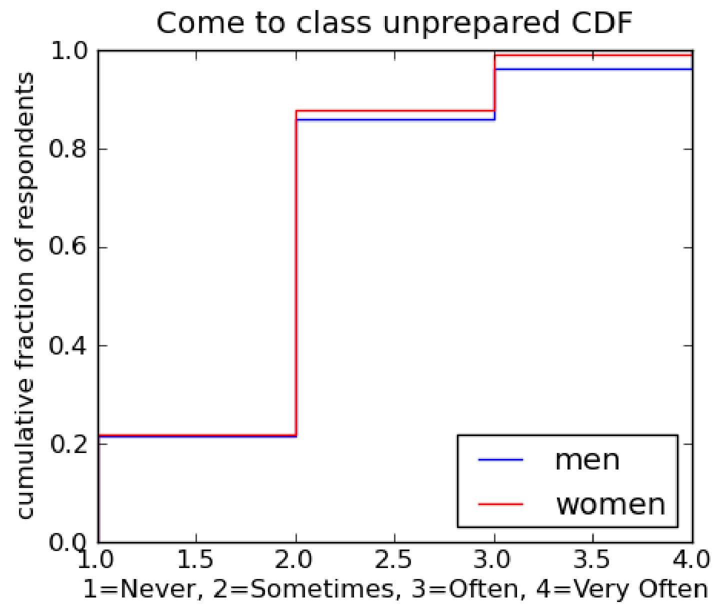

Fig. 5. *This CDF shows results for the survey prompt "Come to class without completing readings or assignments."*

an exponential distribution. An exponential CDF has the equation $CDF(x) = 1 - e^{-\lambda x}$. If a continuous distribution is an effective model for the data, plotting the complementary CDF (CCDF) on a log-y scale should show a straight line with slope $-\lambda$. This CCDF is the right-most graph in Figure 6. The CCDF on a log-y scale looks more like a straight line than not, though it would be easier to see how well the data fits a continuous distribution if there were more possible responses on the survey.

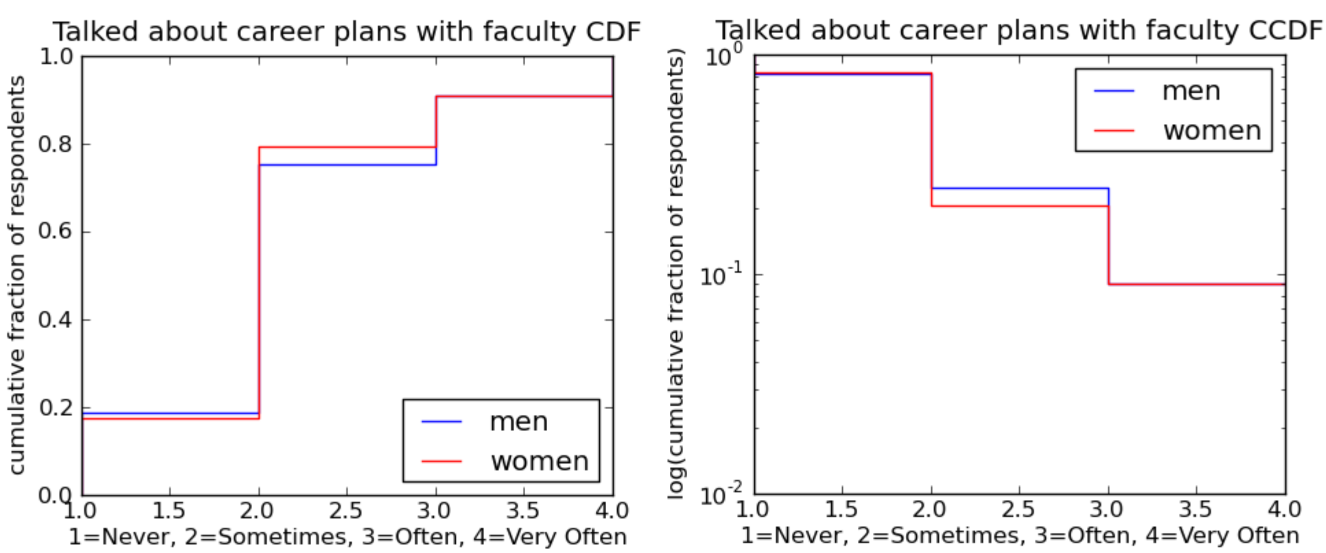

Fig. 6. *This CDF shows results for the survey prompt "Talked about career plans with a faculty member or advisor."*

VIII. Normal Probability Plot

One way to test whether a distribution is normal is to generate a normal probability plot. A quick, simple way to do this is to generate a sample of the same length as a given sample from a normal distribution with μ=0 and σ=1, then plot the data sample against the normal sample. If the sample is normal, then the normal probability plot will be a straight line. For the prompt "Prepared two or more drafts of a paper or assigment before turning it in," the responses of both women and men shown in Figure 7 look like they could be approximated as a normal distribution, so this test is appropriate. The responses of women look more like a normal distribution than the responses of men, so it would make sense if the normal probability plot looked more like a straight line for women than for men. However, since there are only four possible responses, it is difficult to see any significant difference between the normal probability plots for female and male students.

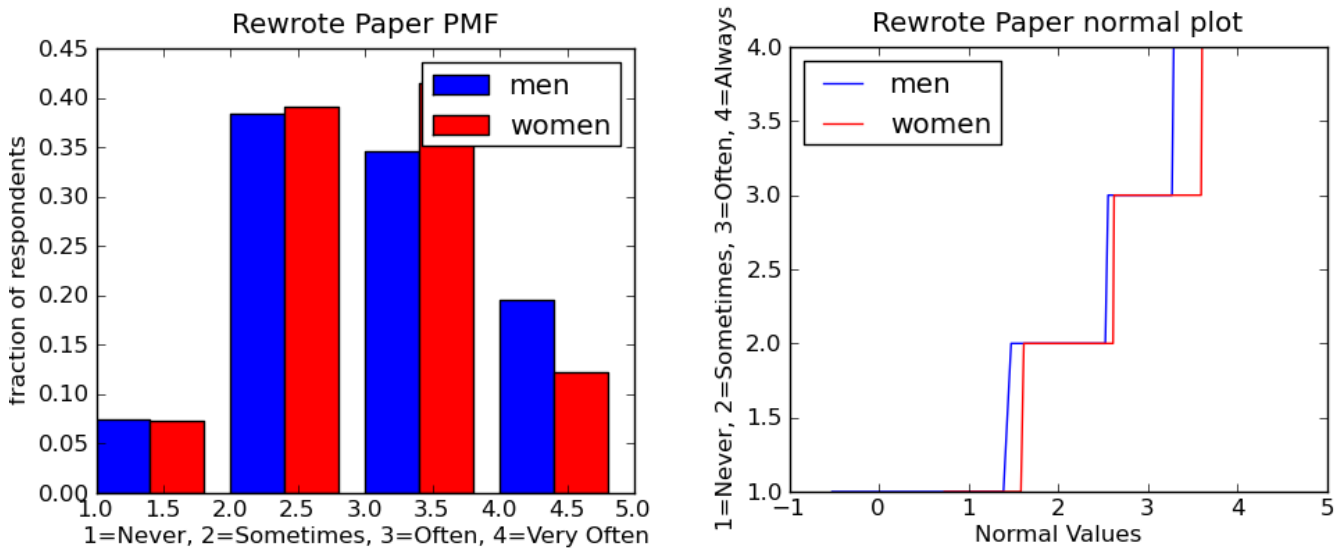

Fig. 7. *This PMF and normal plot show that results from the prompt "Prepared two or more drafts of a paper or assignment before turning it in" could be approximated by a normal distribution.*

IX. Estimating Parameters for the Complete Distribution

Since the data for the prompt "Talked about career plans with a faculty member or advisor" has a CDF that looks exponential, we can estimate the parameter λ for the exponential distribution from which it is a sample, which would be the responses for all first-year students at Olin College from 2007-2010. This simply involves finding the slope of the CCDF plotted on a log-y scale, which is shown in the right-most graph of Figure 6. The slope in this case for both women and men is approximately -0.24, so λ for an exponential model of the complete distribution can be approximated as 0.24. Continuous models like this help approximate what survey responses would be for the entire population of Olin College first-year students.

Since the data for the prompt "Prepared two or more drafts of a paper or assigment before turning it in" shown in Figure 7 can be approximated by a normal distribution, we can estimate parameters for a normal model for the complete distribution. The sample mean for women is 2.59 with a standard deviation of 0.80, and the sample mean for men is 2.63 with a standard deviation of 0.84. Since these data do not have outliers, the sample mean and standard deviation are a reasonable approximation for the mean and standard deviation of the complete population of first-year Olin College students over the time period when the survey was administered.

X. Bayes's Theorem

One question that arises from analyzing PMFs of NSSE data is this: if a student asks questions very often in class, what is the probability that the student is a woman? The answer for the total population of Olin first-year students can be approximated by using Bayes's theorem on data from NSSE. Bayes's theorem represents the relationship between the conditional probabilities of two events. If W is the hypothesis that the student is a woman, E is the fact that the student reported asking questions very often in class, and M is the hypothesis that the student is a man, then the probability that the student is a woman according to Bayes's theorem is as follows:

$$P(W|E) = \frac{P(W)P(E|W)}{P(E)}$$
$$P(E) = P(W)P(E|W) + P(M)P(E|M) \tag{1}$$

Substituting in values from the NSSE data, the probability that a student who asks questions very often is female equals 0.42. This probability is significantly less than the probability that the student is male, which indicates that this gap between genders may apply to the larger population of Olin College first-year students.

XI. Difference-in-Mean Hypotheses

One way to test the significance of the difference between data for women and men is to take two random samples from the pooled data many times and calculate how often the difference in means of the random samples reaches the difference in means of the samples being tested. This fraction of tests is called the p-value of the null hypothesis, where the null hypothesis is that a difference in two data sets is due to chance. This test can be applied to data from female and male respondents to show the relative significance of differences in responses by gender. For the prompt "Ask questions in class or contributed to class discussions," the null hypothesis is that differences in mean responses by gender are due to chance. The p-value of this null hypothesis was 0.001, which indicates that the null hypothesis is highly unlikely and the difference in mean responses by gender is likely significant. For the prompt "Tried to better understand someone else's views by imagining how an idea looks from his or her perspective," the p-value for the same null hypothesis was 0.864, which indicates that the null hypothesis is likely true and the difference in means is likely due to chance. For the prompt "Worked harder than you thought you could to meet an instructor's standards or expectations," the p-value for the same null hypothesis was 0.238, which means the null hypothesis is fairly unlikely but is still possible. It is harder to draw conclusions from p-values that are not very close to 0 or 1, like 0.238.

XII. Chi-Square Hypotheses

Another way to test the significance of the difference between data for women and men is to calculate the difference in the chi squared (χ^2) statistic for two random samples from the pooled data many times and compare it with the difference in χ^2 for the samples being tested. Where O_i is the observed value and E_i is the expected value for each response i,

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i} \tag{2}$$

The p-value for the null hypothesis is analogous to that of a difference-in-mean hypothesis. For the prompt "Ask questions in class or contributed to class discussions," the null hypothesis is that differences in the chi squared statistic by gender are due to chance. The p-value of this null hypothesis was calculated to be 0 because the difference in the chi-square value from random samples of pooled data never reached the sample chi-square value in 10,000 iterations. This indicates that the null hypothesis is highly unlikely and the difference in responses by gender is likely significant. For the prompt "Tried to better understand someone else's views by imagining how an idea looks from his or her perspective," the p-value for the same null hypothesis was 0.0231, which indicates that the null hypothesis is likely false and the difference in means is likely significant. This measure contrasts with the p-value from the difference-in-mean hypothesis for the same prompt, which indicated that the difference between data for women and men was likely due to chance. For the prompt "Worked harder than you thought you could to meet an instructor's standards or expectations," the p-value for the same null hypothesis was 0.0177, which means the null hypothesis is unlikely but is still possible. This p-value is also much smaller than its difference-in-mean analogue, which makes it harder to determine if the difference is significant.

XIII. Fisher's Exact Test

Fisher's Exact Probability Test is yet another way to test the significance of differences between two data samples. It is often used when a sample size is too small for the chi-square test to be appropriate. Figure 8 shows Fisher's Exact Test using the criteria of having a response of 3 or 4 ("Often" or "Very

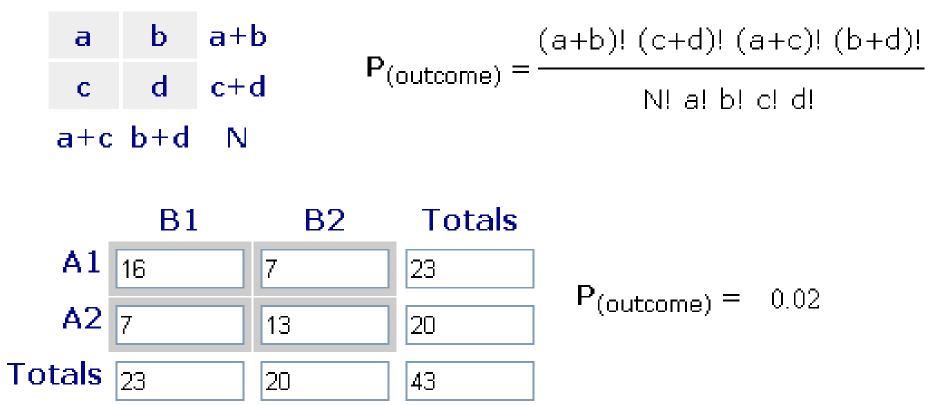

Fig. 8. *Summary of Fisher's Exact Test- images from faculty website for Vassar College.*

Often," respectively) for the prompt "Ask questions in class or contributed to class discussions" as the row headers A1 and A2. Responses of 3 or 4 for the prompt "Talk about career plans with a faculty member or advisor" are the column headers B1 and B2. The null hypothesis was that there is no correlation between the two variables. The test had a p-value of 0.02, which indicates that the null hypothesis is unlikely. This suggests that students who frequently participate in class are more likely to talk about career plans with a faculty member or advisor.

XIV. Pearson's Correlation

Pearson's correlation (ρ) is a measure of how two variables are related. It ranges from -1 to 1, with negative numbers indicating a linear relationship with negative slope (negative correlation), 0 indicating no linear relationship, and positive numbers indicating a linear relationship with positive slope (positive correlation). If μ represents the sample mean, σ represents the sample variance, and n represents the number of data points, Pearson's correlation for two variables x and y can be calculated as follows:

$$p_i = \frac{(x_i - \mu_x)}{\sigma_x}\frac{(y_i - \mu_y)}{\sigma_y}$$
$$\rho = \frac{1}{n-1}\sum_i p_i \tag{3}$$

Pearson's correlation for the responses for the prompts "Ask questions in class or contributed to class discussions" and "Talk about career plans with a faculty member or advisor" is -0.13 for women and 0.24 for men. These numbers indicate that there is a positive correlation between these two survey questions for men but not for women. These relationships are shown in Figure 9.

XV. Linear Regression

Linear regression, which is the process of fitting a line to data, can be done most easily using a linear least squares fit. This method of linear regression minimizes the mean squared error for a linear relationship between variables X and Y with slope α and y-intercept β. The deviation between the data and a linear relationship is equal to $\alpha - \beta x_i - y_i$ for each value in X and Y, and the linear least squares fit minimizes the mean of these deviations. Thus, to calculate the linear least squares fit, with covariance Covar, variance Var, and sample means $\bar{x}$ and $\bar{y}$,

$$\beta = \frac{Covar(X,Y)}{Var(X)}$$
$$\alpha = \bar{y} - \beta\bar{x} \tag{4}$$

A linear regression using this method for the responses of the survey prompt "Ask questions in class or contributed to class discussions" and the means of responses for the prompt "Talk about career plans with a faculty member or advisor" is shown in Figure 9. The means of the responses were used so that there would only be one value corresponding to each response and a linear regression could be performed.

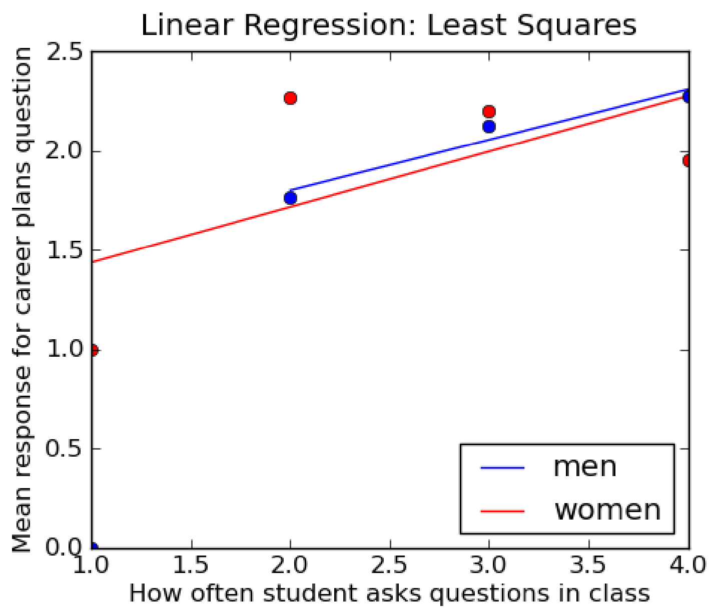

Fig. 9. *A linear regression for the responses of the survey prompt "Ask questions in class or contributed to class discussions" on the x-axis and the means of responses for the prompt "Talk about career plans with a faculty member or advisor" on the y-axis.*

Figure 9 shows that the data for men has a strong positive correlation and the linear least squares method produces an appropriate line of best fit for the data. The Pearson coefficient for this data is 0.28 rather than a value closer to 1, but taking the mean of the responses to the second prompt as was necessary for linear regression produced a stronger correlation than trying to find a linear relationship between the responses for each respondent. The data for women does not seem to have a linear relationship, so the line produced from a linear least squares fit does not represent the data well.

XVI. Conclusions

From these survey responses of Olin College first-year students in the years 2007-2010, two main conclusions can be drawn. Women reported that they "Asked questions in class or contributed to class discussions" less frequently than men, which is demonstrated by the difference in standardized means in Table I, the PMF in Figure 1, the Bayesian analysis in section X, and p-values from differences in mean and chi squared in sections XI and XII. Students who participated in class frequently also reported that they "talked about career plans with a faculty member or advisor" more often that those who did not participate in class as much. This is shown by results from Fisher's Exact Test, Pearson's correlation, and a linear least squares fit in sections XIII-XV. These conclusions are relevant to Olin College in its efforts to provide an equal learning experience for women and men.