

Statistics for Marketing MGT 6469

Sparsh Singhal, Section 2

SPSS Case #5 Vanderbilt Survey Validation

Question 1: Using the VU Student dataset, please answer the following questions:

1. Does the sample accurately reflect the gender make-up of the university? That is, is the sample proportion significantly different from the population proportion?
2. Is the number of men in the sample different from (not equal to) the number of women (mathematically and statistically)?
3. Is the sample significantly different from the equal distribution of Freshmen, Sophomores, Juniors, and Seniors in the Vanderbilt population?
4. In the sample, are year in school and gender independent?

Answer 1:

1.

Gender			
	Observed N	Expected N	Residual
Male	101	119.8	-18.8
Female	149	130.3	18.8
Total	250		

Test Statistics

Gender	
Chi-Square	5.635 ^a
df	1
Asymp. Sig.	.018

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 119.8.

Upon running a chi-square test, we find that - Yes, the sample proportion is significantly different than the population proportion as $X^2_{stat}(1) = 5.635 > X^2_{critical}(1) = 3.841$ and $p = 0.018 < 0.05$.

2.

Gender			
	Observed N	Expected N	Residual
Male	101	125.0	-24.0
Female	149	125.0	24.0
Total	250		

Test Statistics

Gender	
Chi-Square	9.216 ^a
df	1
Asymp. Sig.	.002

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 125.0.

Upon running a chi-square test, we find that - Yes, the number of men in the sample is significantly different than the number of women in the sample as $X^2_{stat}(1) = 9.216 > X^2_{critical}(1) = 3.841$ and $p = 0.002 < 0.05$

3. Frequencies

YearinSchool			
	Observed N	Expected N	Residual
Freshman	62	62.5	-.5
Sophomore	71	62.5	8.5
Junior	52	62.5	-10.5
Senior	65	62.5	2.5
Total	250		

Test Statistics	
YearinSchool	
Chi-Square	3.024 ^a
df	3
Asymp. Sig.	.388

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 62.5.

Upon running a chi-square test, we find that – **NO**, there is no significant difference between the distribution of Freshmen, Sophomores, Juniors, and Seniors in the sample and the Vanderbilt population parameters as $X^2_{stat}(1) = 3.024 < X^2_{critical}(1) = 3.841$ and $p = 0.388 > 0.05$.

4.

Gender * YearinSchool Crosstabulation

			YearinSchool				
			Freshman	Sophomore	Junior	Senior	Total
Gender	Male	Count	24	21	30	26	101
		Expected Count	25.0	28.7	21.0	26.3	101.0
		% within Gender	23.8%	20.8%	29.7%	25.7%	100.0%
		% within YearinSchool	38.7%	29.6%	57.7%	40.0%	40.4%
		% of Total	9.6%	8.4%	12.0%	10.4%	40.4%
	Female	Count	38	50	22	39	149
		Expected Count	37.0	42.3	31.0	38.7	149.0
		% within Gender	25.5%	33.6%	14.8%	26.2%	100.0%
		% within YearinSchool	61.3%	70.4%	42.3%	60.0%	59.6%
		% of Total	15.2%	20.0%	8.8%	15.6%	59.6%
Total	Count	62	71	52	65	250	
	Expected Count	62.0	71.0	52.0	65.0	250.0	
	% within Gender	24.8%	28.4%	20.8%	26.0%	100.0%	
	% within YearinSchool	100.0%	100.0%	100.0%	100.0%	100.0%	
	% of Total	24.8%	28.4%	20.8%	26.0%	100.0%	

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	9.989 ^a	3	.019
Likelihood Ratio	9.967	3	.019
Linear-by-Linear Association	1.185	1	.276
N of Valid Cases	250		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 21.01.

Upon running a chi-square test, we find that – **Yes**, there is a significant difference between the year of school men are in and the year of school females are in. Gender and Year of school are therefore dependent, as $X^2_{stat}(3) = 9.989 < X^2_{critical}(3) = 7.815$ and $p = 0.019 < 0.05$. (See next page for Graph)

Gender and Year in School are Dependent. The year of school men are in and female are in are significantly different.

