

How People Learn, Exercises Week 5 Solution

Question 1

Number of variables	What do you want to do?	What is the nature of the data?	List appropriate tools	Notes on conditions of their use
Working with a single variable	Describing the variable graphically	Nominal	Bar chart, Pie chart	
		Ordinal	Bar chart	
		Interval	Histogram, Box plot, stem and leaf plot	
	Describing the variable in numbers	Nominal	Frequency table, mode	
		Ordinal	Frequency table, median	
		Interval	Mean, median, stem and leaf plot	
	Making inferences about the population	Nominal/ordinal	test of proportion	Random sample (more or less) 2 categories (p and not p) N of at least 20 and at least 10 in each of the 2 categories.
		Interval data sample size of <25	one sample t-test	Random sample (more or less) You believe the population is more or less normally distributed
		Interval data sample size of ≥ 25	One sample z-test	Random sample (more or less)

Question 2

(a)

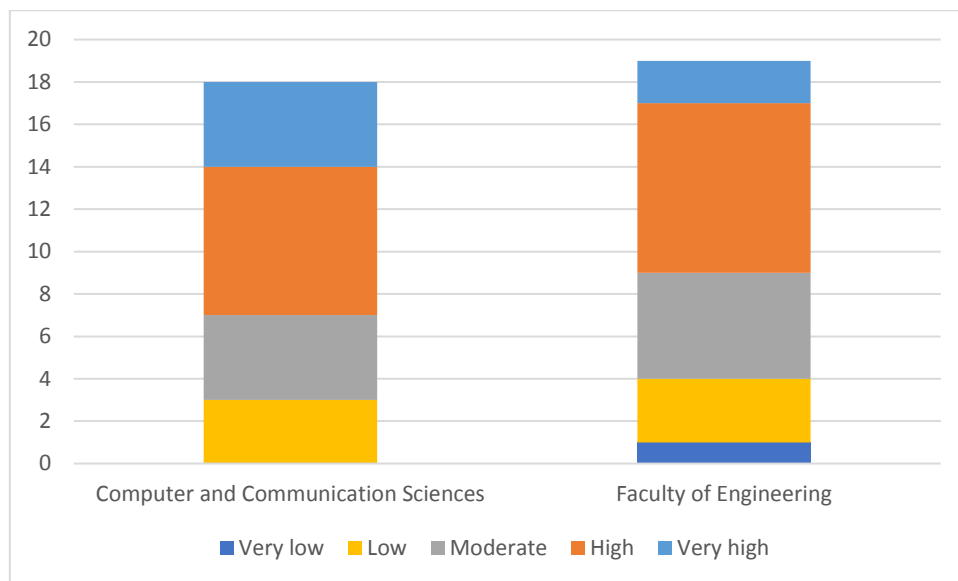
Table 1: Self-monitoring of learning as broken down by faculty of origin

	Computer and Communication Sciences	Faculty of Engineering	Total
Very high	4	2	6
High	7	8	15
Moderate	4	5	9
Low	3	3	6
Very low	0	1	1
	18	19	37

Note: Question is "I often find that I have been reading for class but don't know what it was all about", scores have been reversed (i.e., strongly disagree is presented as very high self-monitoring of learning. N=37, 21 students from other faculties or none not included in this analysis.

(b)

Chart 1: Self-monitoring of learning as broken down by faculty of origin



Note: Question is "I often find that I have been reading for class but don't know what it was all about", scores have been reversed (i.e., strongly disagree is presented as very high self-monitoring of learning. N=41, 12 students from other faculties or none not included in this analysis.

Question 3

(a)

Using the cross tab produced in question 2 it is not possible to do a Chi-square test to test if the two variables ('faculty' and 'score on Q.8') are independent. Why? **The expected values in some of the cells is less than 5.**

(b)

In order to carry out a chi-square test, I need to regroup the data so that all cells have an expected value of at least 5.

Hence I group high scores together and moderate/low scores together.

Hypothesis

H₀: Faculty of origin and self-monitoring score are independent in the wider population

H₁: Faculty of origin and self-monitoring score are not independent in the wider population

What is the critical value for this test?: The table has 2 rows and 2 columns. The degrees of freedom is $(r-1)*(c-1) = (1*1)=1$. The critical value for $p=0.05$ $df=2$ for chi-square is 3.841 (you get this from a Chi-square distribution table such as this: <http://www.statisticshowto.com/tables/chi-squared-table-right-tail/>)

(I have used the reversed Q.8 in this question)

Observed Values

	IC	STI	Total
High scores	11	10	21
Moderate/ Low scores	7	9	16
	18	19	37

Expected Values (column total row total/ overall total)*

	IC	STI	total
High scores	10.22	10.78	21
Moderate/ Low scores	7.78	8.22	16
	18	19	37

Note that all expected values are greater than 5 so we can use Chi square

$(\text{Observed Values} - \text{Expected Values})^2 / \text{Expected Values}$

	IC	SB	STI
High scores	0.060	0.057	
Moderate/ Low scores	0.079	0.075	

Chi-square= 0.271

Note: Excel can calculate chi square p value once you have observed and expected tables. The p value for this is p=0.60. (You can confirm this matches your actual chi-square statistic calculated by hand by looking at how the chi-square statistic is located on the chi-square distribution table <http://www.statisticshowto.com/tables/chi-squared-table-right-tail/>).

Conclusion

The chi-square statistic is a lot less extreme than the critical value. Therefore we do not reject the null hypothesis that faculty of origin and self-monitoring are independent of each other. We have no evidence that there is an association between self-monitoring and faculty of origin in the wider population.