

WEB PAGE FOR CHAPTER 14

MULTIPLE CHOICE QUESTIONS

- 1 The sum of expected frequencies in chi square must equal N:
 - (a) true
 - (b) false
 - (c) depends on the degree of association between the expected and the observed
 - (d) only when the distribution is normal
- 2 The sum of the differences between observed and expected frequencies must be zero:
 - (a) true
 - (b) false
 - (c) only when the distribution is normal
 - (d) sometimes
- 3 As the number of categories in a chi square table increases so does df:
 - (a) true
 - (b) false
 - (c) sometimes
 - (d) depends on N
- 4 As chi square value increases so does the likelihood of rejecting the null hypothesis:
 - (a) true
 - (b) false
 - (c) sometimes
 - (d) depends on df
- 5 On rejecting the null hypothesis, for the chi square test, the researcher should then inspect the pattern of the tabled frequencies and counts to the significant associations:
 - (a) true
 - (b) false
 - (c) sometimes
 - (d) depends on df
- 6 The chi square goodness of fit test assumes:
 - (a) normal distribution against which the observed data will be compared
 - (b) a non-normal distribution against which the observed data will be compared
 - (c) a particular distribution against which the observed data will be compared
 - (d) no known distribution
- 7 Given a 3×4 chi square, how many df are there?
 - (a) 12
 - (b) 11
 - (c) 9
 - (d) 6
- 8 What level of measurement is used with chi square?
 - (a) nominal
 - (b) ordinal
 - (c) interval
 - (d) ratio
- 9 If there are eight categories or classes in a goodness of fit test how many df are there?
 - (a) 8
 - (b) 7
 - (c) 6
 - (d) cannot tell from the information given

- 10 In a goodness of fit test to determine whether there are significantly more petrol (gas) price increases on particular days of the week, what percentage of the total increases should be expected on each day of a seven day week?
- 7%
 - 70%
 - 14.3%
 - 7/10ths

SPSS ACTIVITIES

- Access SPSS Chapter 12 Data File B and:
 - Test the goodness of fit null hypothesis are not significantly biased towards any particular qualification.
 - Test the goodness of fit null hypothesis that smokers and non-smokers are equally represented in the sample.
 - Test the null hypothesis that there is no significant association between whether a person smokes or not and their drink preference.
 - Test the null hypothesis that there is no statistical association between gender and qualifications held.

Interpret your printout for each and discuss the results in class.

ADDITIONAL QUESTIONS (For group work – discuss your answer in your groups)

- In groups, inspect the following tables relating to staff in different departments in a company and answer the following questions:
 - Why are the expected frequencies all the same at 138.3?
 - What is the null hypothesis?
 - Is chi square significant?
 - What are the implications of these results for the null hypothesis?
 - Why is $df = 2$?

Chi-Square Test Frequencies

DEPART			
	Observed N	Expected N	Residual
Sales	271	138.3	132.7
Financial services	67	138.3	-71.3
producton	77	138.3	-61.3
Total	415		

Test Statistics	
	DEPART
Chi-Square ^a	191.210
df	2
Asymp. Sig.	.000

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 138.3.

- In your group, inspect the following tables and answer the following questions:
 - The number of people in this analysis is?
 - Express the null hypothesis.
 - Is the null hypothesis rejected? Explain your answer using information from both tables.
 - Why is $df = 2$?

Crosstabs

GENDER * DEPART Crosstabulation

			DEPART			Total
			Sales	Financial services	Production	
GENDER	Female	Count	116	28	19	163
		Expected Count	105.8	27.0	30.2	163.0
		% within GENDER	71.2%	17.2%	11.7%	100.0%
		% within DEPART	44.1%	41.8%	25.3%	40.2%
		% of Total	28.6%	6.9%	4.7%	40.2%
Male	Count	Count	147	39	56	242
		Expected Count	157.2	40.0	44.8	242.0
		% within GENDER	60.7%	16.1%	23.1%	100.0%
		% within DEPART	55.9%	58.2%	74.7%	59.8%
		% of Total	36.3%	9.6%	13.8%	59.8%
Total		Count	263	67	75	405
		Expected Count	263.0	67.0	75.0	405.0
		% within GENDER	64.9%	16.5%	18.5%	100.0%
		% within DEPART	100.0%	100.0%	100.0%	100.0%
		% of Total	64.9%	16.5%	18.5%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	8.632 ^a	2	.013
Likelihood Ratio	9.044	2	.011
Linear-by-Linear Association	7.531	1	.006
N of Valid Cases	405		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 26.97.

- 3 Study the SPSS output data given in the tables below in groups then answer questions 3 – 9 following below.

Table 1: District * Contract Status Crosstabulation

			Contract Status		
			Competitive Contract	Fixed Contract	Total
District	Other District	Count	114	25	139
		% within District	82.0%	18.0%	100.0%
		% within Contract Status	61.6%	50.0%	59.1%
		% of Total	48.5%	10.6%	59.1%
	South Florida District	Count	71	25	96
		% within District	74.0%	26.0%	100.0%
		% within Contract Status	38.4%	50.0%	40.9%
		% of Total	30.2%	10.6%	40.9%
	Total	Count	185	50	235
		% within District	78.7%	21.3%	100.0%
		% within Contract Status	100.0%	100.0%	100.0%
		% of Total	78.7%	21.3%	100.0%

Table 2: Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2.200(b)	1	.138		
Continuity Correction(a)	1.745	1	.186		
Likelihood Ratio	2.174	1	.140		
Fisher's Exact Test				.148	.094
Linear-by-Linear Association	2.191	1	.139		
N of Valid Cases	235				

a Computed only for a 2x2 table

b 0 cells (.0%) have expected count less than 5. The minimum expected count is 20.43.

- 3 What two variables are involved in this study?
 - (a) Pearson and chi-square
 - (b) count and total
 - (c) total and contract status
 - (d) contract status and district
 - (e) observed and expected

- 4 Table 1 above is generally called a(n) table.
 - (a) contiguous
 - (b) contagious
 - (c) contingency
 - (d) output
 - (e) kitchen

- 5 What type of test is being used above?
 - (a) Goodness of fit
 - (b) Comparison of percentages
 - (c) Comparison of means
 - (d) Statistical
 - (e) Crosstabs

- 6 Write a null hypothesis for the study.
- 7 What is the obtained p value?
 - (a) .148
 - (b) .138
 - (c) 2.2
 - (d) .094
 - (e) 1
- 8 What conclusion could you draw at the .05 level?
 - (a) the null hypothesis is retained
 - (b) the alternative hypothesis is not rejected
 - (c) the null hypothesis is not supported
 - (d) the hypothesis is rejected
 - (e) the hypothesis is null and void
- 9 Both variables are measured on what type of scale?
 - (a) ordinal
 - (b) nominal
 - (c) interval
 - (d) ratio
 - (e) normal

CLASS ACTIVITIES AND DISCUSSIONS

- 1 Collect some nominal data from the class to test against a chi square goodness of fit; for example, preferences for pop singers, football teams, political parties, etc. Place the data on SPSS. Interpret your result.
- 2 Using the data you have collected above, divide your sample using some classification such as gender, ethnic group, possession of satellite TV, etc. and perform a chi square of association (crosstabs). Interpret your result.

ANSWERS TO MULTIPLE CHOICE QUESTIONS

- 1 (a), 2 (a), 3 (a), 4 (d), 5 (a), 6 (c), 7 (d), 8 (a), 9 (b), 10 (c).

Web Page for Chapter 15

MULTIPLE CHOICE QUESTIONS

- 1 Why is it important to plot a scattergraph before calculating a correlation coefficient for a large data set?
 - (a) to check whether both variables are normally distributed
 - (b) to find out whether the correlation is positive or negative
 - (c) to ensure the plots are roughly linear
 - (d) to ensure the data are all present
- 2 A student produces a correlation of +1.3. This is:
 - (a) a high positive correlation
 - (b) a significant correlation
 - (c) an impossible correlation
 - (d) only possible if N is large
- 3 Which statement is true? Correlations
 - (a) reveal significant differences
 - (b) do not imply cause and effect
 - (c) require nominal data only
 - (d) cannot be used with ordinal data

- 4 The Spearman correlation is used with:
 - (a) ordinal data
 - (b) interval data
 - (c) nominal data
 - (d) ratio data
- 5 When you are correlating two interval variables you can use:
 - (a) the 'r' correlation
 - (b) the Pearson coefficient
 - (c) the Spearman correlation
 - (d) the coefficient of determination
- 6 What sort of correlation would be expected between a company's expenditure on health and safety and the number of work related accidents?
 - (a) positive
 - (b) negative
 - (c) random
 - (d) none
- 7 If Ali scored the top mark in the apprentices test on computing and the correlation between that test and the test on finger dexterity was +1.0, what position did Ali get in the second test?
 - (a) middle
 - (b) bottom
 - (c) top
 - (d) cannot say from the information given
- 8 Which correlation is the strongest, +.65 or $-.70$?
 - (a) $-.70$
 - (b) +.65
 - (c) depends on N
 - (d) cannot say from the information given
- 9 The factor that affects the significance of a correlation most is:
 - (a) whether both variables are interval
 - (b) the p value
 - (c) df
 - (d) all of the above
- 10 Kendall's tau is used for correlating:
 - (a) true dichotomies
 - (b) artificial dichotomies
 - (c) ranked data
 - (d) continuous scores
- 11 All correlation coefficients share in common the property that they range between:
 - (a) +1 and 0
 - (b) +1.00 and -1.00 .
 - (c) +.1 and $-.1$
 - (d) +1.96 and -1.96
- 12 The process to rule out the influence of one or more variables upon the criterion variable is:
 - (a) regression
 - (b) partial correlation
 - (c) multiple correlation
 - (d) Yate's correction

- 13 A researcher is interested in the relationship between motivation and academic performance among MBA students. However, a colleague warns her that IQ may influence both variables. How can she deal with this situation?
 - (a) use zero order correlations
 - (b) inspect a scattergraph of motivation versus academic performance
 - (c) use ANOVA
 - (d) use the partial correlation technique
- 14 The symbol for the correlation co-efficient is:
 - (a) p
 - (b) Σ
 - (c) σ
 - (d) α
 - (e) r
- 15 In a study relating mean house mortgage interest rates per annum and the number of new applicants for mortgages over the last 10 years:
 - (a) the mean house mortgage interest rate per annum is the dependent variable
 - (b) in a scattergraph, the mean house mortgage interest rate per annum would be placed on the Y axis
 - (c) the number of new applicants per annum is represented on the Y axis
 - (d) in a scattergraph, the number of new applicants per annum is the independent variable represented on the Y axis

SPSS ACTIVITIES

Access Chapter 13 Data File A

- 1 Using Pearson's technique, obtain the correlations between (a) current salary and starting salary, (b) self concept score and anxiety score, and (c) absence and self-concept. Interpret your findings.
- 2 Using Spearman's rho, obtain correlations between (a) pretest and post-test scores, (b) absence and salary now, and (c) performance after training and current salary. Remember, you must use ordinal data. Interpret your findings.
- 3 Produce scattergraphs for the pairings (a) absence with self-concept, and (b) self-concept and anxiety.
- 4 Carry out a partial correlation between (a) self-concept and anxiety, holding age constant, and (b) train1 and train2, holding experience constant. What do the results imply?

MULTIPLE CHOICE QUESTIONS – SET B

- 1 The co-efficient of determination ranges in value from:
 - (a) 0 to 1
 - (b) -100 to +100
 - (c) 0 to infinity
 - (d) 0% to 100%
 - (e) -100% to +100%
- 2 In a correlation study of the relationship between profits and sales, SPSS outputs show that the correlation co-efficient equals +.9 with a significance level of $p < .001$. H_1 : The level of profits is significantly related to the level of sales.

The correlation co-efficient is:

- (a) strong, positive and significant
- (b) moderate, positive and not significant
- (c) weak, negative and not significant
- (d) strong, positive and not significant
- (e) indeterminate

- 3 With reference to (2) above the hypothesis as stated is:
- (a) not rejected
 - (b) accepted
 - (c) supported
 - (d) not supported
 - (e) all of the above
- 4 When 'r' is negative, one variable increases in value:
- (a) the other increases
 - (b) the other increases at a greater rate
 - (c) the other variable decreases in value
 - (d) there is no change in the other variable
 - (e) all of the above
- 5 For a normal good, if price increases then the quantity demanded decreases. What type of correlation co-efficient would you expect in this situation?
- (a) 0
 - (b) positive
 - (c) .9
 - (d) negative
 - (e) unknowable
- 6 If the probability of there being no significant relationship between two variables in a correlation test equals .02, then we can conclude:
- (a) there is not a statistically significant relationship between the variables at the .10 level
 - (b) there is a statistically significant relationship between the variables at the .01 level
 - (c) there is not a statistically significant difference between the variables at the .05 level
 - (d) there is a statistically significant relationship between the variables at the .05 level
- 7 The Co-efficient of Determination expresses the amount of the variation in the dependent variable that can be explained by:
- (a) the proportion of Y
 - (b) the variation in the independent variable Y
 - (c) the variation in the independent variable X
 - (d) all of the above
 - (e) none of the above
- 8 To obtain the Co-efficient of Determination, it is necessary to:
- (a) take the square root of 'r'
 - (b) compute the exponential value of 'r'
 - (c) square 'r'
 - (d) $r \times r \times r$
 - (e) take the cube root of 'r'
- 9 If $r = .7$, then the Co-efficient of Determination is equal to:
- (a) 7.0
 - (b) .7
 - (c) .049
 - (d) .49
 - (e) 4.9%
- 10 If two variables are absolutely independent of each other the correlation between them must be:
- (a) -1
 - (b) 0
 - (c) +1
 - (d) +.1

ADDITIONAL QUESTIONS/PROBLEMS FOR GROUP WORK

- 1 The annual national beauty contest is under way. Each of the 10 contestants receives a rank on both personality and appearance. State a relevant hypothesis. The correlation is $+0.03$. Discuss as a group which correlation coefficient would have been used. What conclusions do you draw about personality and appearance?
- 2 A researcher has tested this hypothesis: *'There is a negative relationship between the finger dexterity scores and the data input error rate'*. Does the hypothesis stated require a one-tailed or a two-tailed test of significance? State the null hypothesis being tested. Which correlation test would be used? If the result is significant at the 5% level state what conclusion is made about the null hypothesis?
- 3 If leadership and productivity are perfectly positively correlated, can you conclude that to improve productivity you should improve leadership? If not, why not?
- 4 Over the 365 days in the year, the greater the quantity of ice cream sold the greater the number of murders.
 - (a) what is the direction of this correlation?
 - (b) what does this correlation imply?
- 5 Match the statistic with the definition. (There is one definition without a statistic provided):

(a) partial correlation	(i) used when both variables can be split into artificial dichotomies at critical points
(b) Spearman's rho	(ii) both variables must be continuous
(c) Product moment correlation (Pearson)	(iii) used with ranks and also when number of cases is under 30
	(iv) used to control a variable that might mask the correlation between two other variables
- 6 A researcher correlates personal income with attendance at religious services over one year. The correlation is $+0.57$. But when the researcher controls for the effects of age, the partial correlation falls to $+0.14$. Can you explain what the researcher did and why? What does the result imply? Why has the size of the correlation fallen so much?

CLASS ACTIVITIES

- 1 Carry out a class activity by requesting everyone to indicate how many hours they spent studying the previous chapter on chi square and the number of items they got correct on the questions set by the teacher/lecturer. Correlate these two sets of data and see what they reveal.
- 2 Study the following correlation matrix then answer the questions below.

Correlations

		Work interferes with my social life	I spend too much time travelling to work	Personal issues reduce my work effectiveness	I struggle with living expenses
Work interferes with my social life	Pearson Correlation	1	.436(**)	.159	.026
	Sig. (2-tailed)	.	.000	.158	.817
	N	80	80	80	80
I spend too much time travelling to work	Pearson Correlation	.436(**)	1	.223(*)	-.013
	Sig. (2-tailed)	.000	.	.047	.907
	N	80	80	80	80
Personal issues reduce my work effectiveness	Pearson Correlation	.159	.223(*)	1	.235(*)
	Sig. (2-tailed)	.158	.047	.	.036
	N	80	80	80	80
I struggle with living expenses	Pearson Correlation	.026	-.013	.235(*)	1
	Sig. (2-tailed)	.817	.907	.036	.
	N	80	80	80	80

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

- Identify variables that are positively correlated. At what level of significance are they?
- Which level of data has been used?
- How many subjects were there?
- State a null hypothesis linking work inference with social life and struggling with living expenses. Is this null hypothesis rejected or supported? Explain your answer.
- What does the correlation between 'I struggle with living expenses' and 'I spend too much time travelling to work' imply?

ANSWERS TO MULTIPLE CHOICE QUESTIONS – SET A

1 (c), 2 (c), 3 (b), 4 (a), 5 (b), 6 (b), 7 (c), 8 (a), 9 (c), 10 (b), 11 (b), 12 (b), 13 (d), 14 (e), 15 (c).

ANSWERS TO MULTIPLE CHOICE QUESTIONS – SET B

1 (d), 2 (a), 3 (c), 4 (c), 5 (d), 6 (d) 7 (c), 8 (c), 9 (d), 10 (b)

ANSWERS TO QUESTIONS IN TEXT

Qu. 15.1 (a) +, (b) +, (c) +, (d) –, (e) +, (f) –.

Qu. 15.2 –1

Qu. 15.3 (b)

Qu. 15.4 bottom

Qu. 15.5 (d)

Qu. 15.6 75% unexplained

Qu. 15.7 (b)