

SO 2010

UNIVERSITY OF WARWICK

Summer Examinations 1995

SURVEYS, SECONDARY ANALYSIS, AND SOCIAL STATISTICS

Candidates should answer THREE questions, including at least ONE from Section A and at least ONE from Section B. In Section A candidates are required to provide commentaries on their answers.

Time allowed: 2 hours

Read carefully the instructions on your answerbook and make sure that the particulars required are entered on each answerbook.

Approved calculators may be used

SECTION A

- 1 The mean number of months unemployed in the last five years for a random sample of 576 British men aged 50-59 was found to be 6.73 months, with a sample standard deviation of 12.00 months.
 - (i) Calculate a 95% confidence interval for the mean number of months unemployed in the last five years for men aged 50-59.
 - (ii) The mean number of months unemployed in the last five years for all men of working age is known to be 6.33 months. Calculate a z-statistic and use it to test whether this is a plausible mean number of months unemployed in the last five years for men aged 50-59.
 - (iii) If the population standard deviation for the mean number of months unemployed in the last five years for women aged 50-59 is assumed to be 9.00, how big a sample would be needed to produce a sample mean that one could be 95% confident fell within one month of the population mean number of months unemployed in the last five years for women aged 50-59? Comment on your answer in relation to your answer to part (i) of this question.

Explain how and why your answer to (i) enables you to answer part (ii) without calculating a z-statistic.

/continued.....

-
- 2 The following table shows whether men had degrees according to their 'race' for a (stratified) random sample of 300 British 30-year old men.

<u>Degree</u>	<u>Yes</u>	<u>No</u>	<u>TOTAL</u>
White	18	102	120
'Black'	6	54	60
'Asian'	12	48	60
Chinese	24	36	60
TOTAL	60 (20%)	240 (80%)	300

'Black' = Black Caribbean; Black African; Black Other.

'Asian' = Indian; Pakistani; Bangladeshi.

- (i) Calculate the chi-square statistic for the above table and use it to test the hypothesis that there is no relationship between having a degree and 'race' for 30-year old men.

(Note: the critical value at the 5% level of a chi-square statistic with 3 degrees of freedom is 7.81).

- (ii) A similar table based on a similarly stratified random sample of 300 30-year old women, once again showing the relationship between having a degree and 'race', gave rise to a chi-square statistic of 0.63. Use Cramer's V to compare the strengths of the relationships in the two tables, and explain why the two chi-square statistics could also have been used for this purpose.

- (iii) Does the above table suggest that the likelihood of having a degree differs between white, 'Black' and 'Asian' 30 year-old men?

(Note: the critical value at the 5% level of a chi-square statistic with 2 degrees of freedom is 5.99).

/continued.....

- 3 The following table shows the mean values of an index of how many household tasks married women feel they should perform according to their highest educational qualification in a random sample of 774 British women.

<u>Qualification</u>	<u>Mean</u>	<u>s</u>	<u>n</u>
Degree	5.70	1.10	48
`A' level	6.20	1.34	144
`O' level	6.42	1.53	166
None of above	6.81	1.52	416
		TOTAL	774

(s is sample standard deviation; n is sample size).

- (i) Test the hypothesis that, in the population, all four groups of women have the same mean value for the index. Discuss your findings with reference to the sample means.

(Note: the critical value of F at the 5% level corresponding to 3 degrees of freedom and 770 degrees of freedom is 2.61; the between-groups and within-groups sums of squares are 108.0 and 1848.0 respectively).

- (ii) Test the hypothesis that, in the population, women whose highest qualifications are degrees and women whose highest qualifications are `A' levels have the same mean value for the index.

(Note: the critical value of t at the 5% level corresponding to 190 degrees of freedom is 1.97; the pooled sample standard deviation for women whose highest qualifications are degrees and women whose highest qualifications are `A' levels is 1.28)

/continued.....

- 4 In a random sample of 34 countries, the (Pearson) correlation between total fertility rate (children per female lifetime) and per capita GNP (Gross National Product; measured in thousands of US dollars) was found to be -0.60.

- (i) Test the hypothesis that there is no relationship between total fertility rate and per capita GNP.

(Note: the critical value of F at the 5% level corresponding to 1 degree of freedom and 32 degrees of freedom is 4.15).

The regression equation corresponding to the dependence of total fertility rate on per capita GNP is

$$y = 4.83 - 0.22x_1$$

where y is total fertility rate (children per female lifetime), and x_1 is per capita GNP (in thousands of US dollars).

- (ii) Use the above equation to predict the total fertility rate in Switzerland (per capita GNP = 20 thousand US dollars). Does the prediction seem sensible? The actual total fertility rate for Senegal is 6.40 children per female lifetime. What sort of per capita GNP would be needed for the equation to predict this fertility rate? Why might the above linear regression equation be an inappropriate model of the relationship between total fertility rate and per capita GNP?

The addition to the regression analysis of a second independent variable, x_2 , which is the percentage of married women using modern contraception, results in the following equation

$$y = 6.45 - 0.04x_1 - 0.07x_2$$

- (iii) Explain why the coefficient of x_1 , per capita GNP, changes between the two equations.

/continued.....

5 The following crosstabulation is of party identification [P] by social class [C] by region [R] of England for a random sample of 1,388 electors. (Electors who identified with a party other than Conservative or Labour have been excluded):

REGION = London	<u>Middle class</u>	<u>Working class</u>	<u>TOTAL</u>
Conservative	52	33	85
Labour	<u>24</u>	<u>65</u>	<u>89</u>
TOTAL	76	98	174

REGION = South	<u>Middle class</u>	<u>Working class</u>	<u>TOTAL</u>
Conservative	173	153	326
Labour	<u>42</u>	<u>100</u>	<u>142</u>
TOTAL	215	253	468

REGION = North	<u>Middle class</u>	<u>Working class</u>	<u>TOTAL</u>
Conservative	161	199	360
Labour	<u>68</u>	<u>318</u>	<u>386</u>
TOTAL	229	517	746

- (i) Use odds ratios to summarise the way in which the relationship between party identification and class varies according to region. The chi-square statistics for the three sub-tables are 20.7, 22.0 and 64.3. Using these chi-square statistics, test the relationship in each sub-table for significance.
- (ii) Use odds ratios to summarise the relationships between party identification and region, and between class and region.
- (iii) Use the following results corresponding to the fit of various log-linear models to determine the most appropriate model of the table given above. Justify your choice, and, given the model that you have selected, comment on your findings in parts (i) and (ii)

(Note: the critical value at the 5% level of a chi-square statistic with 2 degrees of freedom is 5.99; the critical value at the 5% level of a chi-square statistic with 1 degree of freedom is 3.84).

Model No.	Model	Deviance	d.f.	P	Change in deviance	d.f.	P	Compared to model
1.	[P] [C] [R]	199.3	7	0.000				
2.	[PC] [R]	77.9	6	0.000	121.4	1	0.000	1
3.	[CR] [P]	167.5	5	0.000	31.8	2	0.000	1
4.	[PR] [C]	141.1	5	0.000	58.2	2	0.000	1
5.	[PC] [CR]	46.1	4	0.000	31.8	2	0.000	2
6.	[PC] [PR]	19.7	4	0.001	58.2	2	0.000	2
7.	[PR] [CR]	109.3	3	0.000	31.8	2	0.000	4
8.	[PC] [PR] [CR]	2.0	2	0.362	17.7	2	0.000	6
9.	[PCR]	0.0	0	-----	2.0	2	0.362	8

/continued....

SECTION B

- 6 Discuss the strengths and weaknesses of the data collected by ONE existing social survey, focussing on the potential of the data as a resource for secondary analysts.
-
- 7 'The production of a good questionnaire is as much dependent on the valid operationalisation of concepts as on the sensible wording of questions.' Discuss, with reference to ONE real or hypothetical survey-based research project.
-
- 8 'Social survey interviews and ethnographic interviews have more in common than is generally acknowledged.' Discuss.
-
- 9 'Sociologists have no incentive to use qualitative and quantitative methods in combination, since, even if criticisms of quantitative methods are ignored, the data produced by the two approaches are epistemologically unreconcilable.' Discuss.
-
- 10 Critically discuss the following table. Your discussion should include:
- * an account of what you would like to know about the data collection process and the sample;
 - * a consideration of the validity of the variables as indicators of underlying concepts;
 - * a description of the substantive relationship visible in the table;
 - * an outline of how the analysis needs to be extended and/or could be elaborated.

[Note: You may assume that the overall relationship in the table is statistically significant; you should specify any more focussed statistical tests that you would ideally like to carry out].

RELIGION by ATTITUDE TOWARDS CHILDREN

<u>Religion</u>	<u>Attitude towards children</u>					
	Score = Low		Score = Medium		Score = High	
		%		%		%
Church of England	128	(29.3)	155	(35.5)	154	(35.2)
Roman Catholic	34	(26.2)	41	(31.5)	55	(42.3)
Other Christian	42	(24.9)	66	(39.1)	61	(36.1)
Non-Christian	9	(40.9)	7	(31.8)	6	(27.3)
No religion	157	(37.3)	164	(39.0)	100	(23.8)

[Note: The variable 'Attitude towards children' is a scale. Low scores on the scale supposedly correspond to a 'pragmatic' attitude towards children, i.e. the worth of having children is evaluated by balancing costs and benefits to the parents; high scores on the scale supposedly correspond to a 'fundamentalist' attitude towards children, i.e. one that sees having children as being an inherently valuable activity].