

SO5041 Assignment 3, Autumn 2017

November 2017

Deadline

Friday 1 Dec 2017 4pm, by Sulis

1 Comparing means

(Marks: 10)

Download this small Stata dataset from

<http://teaching.sociology.ul.ie/so5041/assign3.dta>.

Use Stata to compare the mean income of men with that of women. Interpret the results in terms of the null hypothesis that mean income is the same in the two groups.

2 European Social Survey

(Marks: 20)

The European Social Survey is carried out approximately every two years in a large number of European countries. Among the questions it asks is the following:

'Please say to what extent you agree or disagree with each of the following statements. Gay men and lesbians should be free to live their own life as they wish'

This is answered on a scale from 1, "Strongly agree", to 5, "Strongly disagree".

The following table (Table 1) and graphs (Figs 1 and 2) summarise how the answers differ by survey wave, gender and country, for a selection of western European countries.

- Use the tables and graphs to write a short report summarising how attitudes differ by gender and country, and how this changes over time (15 marks)
- Use the table and its reported chi-square to conduct a statistical test for country differences in attitude (5 marks)

Table 1: Attitudes to homosexuality, five western European countries, ESS 2016

. tab centry freehms if wave==2016, chi row

Key
<i>frequency</i>
<i>row percentage</i>

Country	Gays and lesbians free to live life as they wish					Total
	Agree str	Agree	Neither a	Disagree	Disagree	
DE	1,261 44.45	1,219 42.97	193 6.80	103 3.63	61 2.15	2,837 100.00
FR	1,338 65.27	467 22.78	118 5.76	56 2.73	71 3.46	2,050 100.00
GB	841 43.85	831 43.33	153 7.98	63 3.28	30 1.56	1,918 100.00
IE	1,101 41.05	1,194 44.52	195 7.27	134 5.00	58 2.16	2,682 100.00
SE	984 64.15	450 29.34	69 4.50	17 1.11	14 0.91	1,534 100.00
Total	5,525 50.13	4,161 37.76	728 6.61	373 3.38	234 2.12	11,021 100.00

Pearson chi2(16) = 548.5128 Pr = 0.000

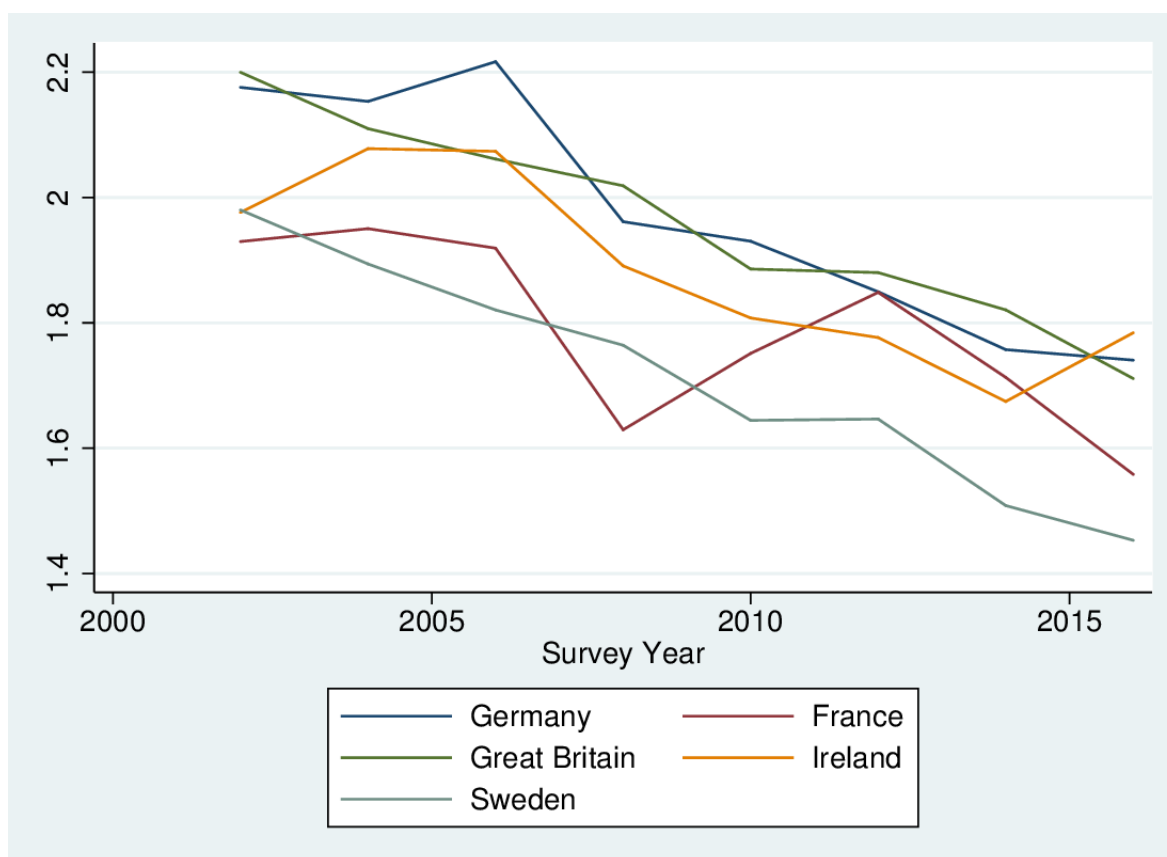


Figure 1: Attitudes to homosexuality, ESS, by country and year (average score on 1-5 scale)

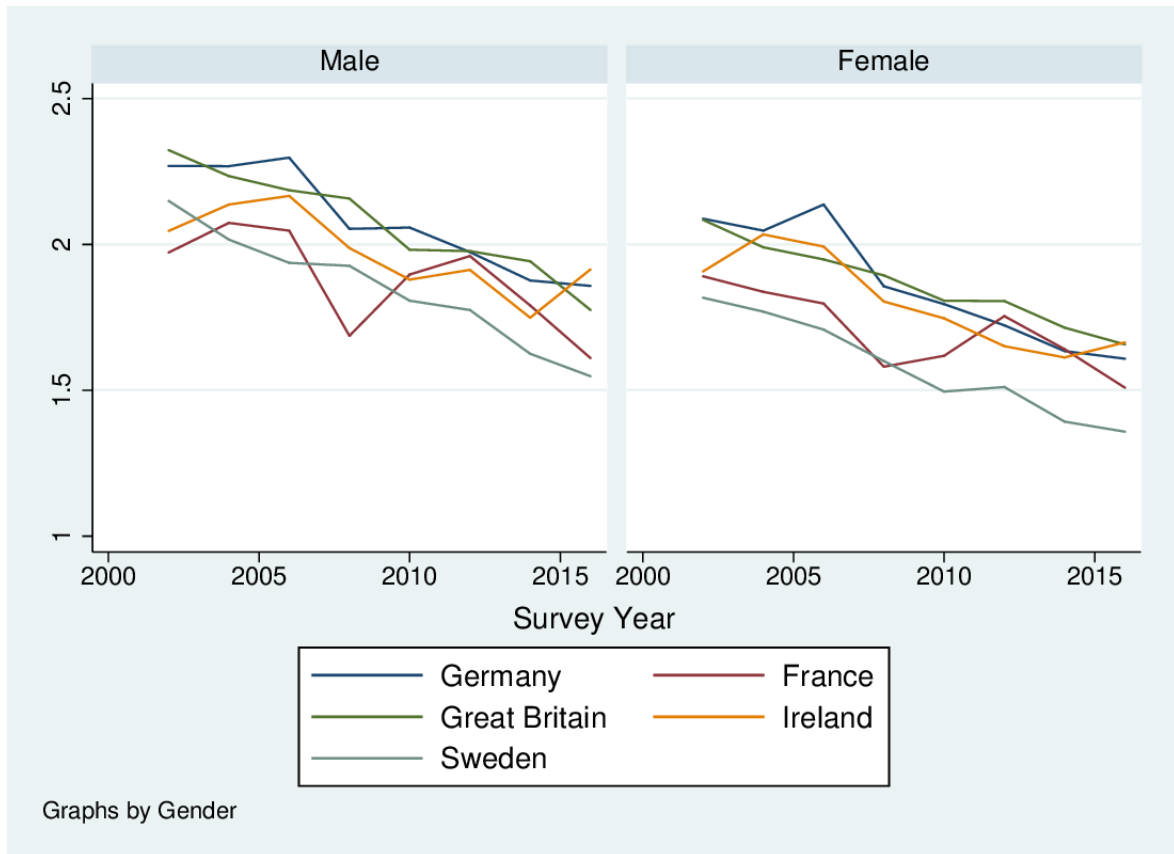


Figure 2: Attitudes to homosexuality, ESS, by country, gender and year (average score on 1-5 scale)

3 Work and TV

(Marks: 10)

A sample of adults has been asked to record how much time they spend watching television in a given week. They have also been asked to record their work hours. Researchers feel that time spent working will affect time spent watching television.

- Do you think this is a plausible idea? If so, how would the connection work and what direction do you think it would take?

To test this hypothesis, they conduct a regression analysis, with *TV* time in minutes per day as the dependent variable and *Work* time (in hours per week) as the explanatory variable. That is, $TV = a + b \times Work$. They get the following results:

- Constant: 37
- Coefficient for hours work: -0.5
- Draw this line on a graph, with *TV* time on the Y-axis, and *Work* time on the X-axis
- If someone works exactly 40 hours, what is his/her predicted time watching television?
- If someone works 15 hours, how much television do they watch?
- If someone works 168 hours, how much television do they watch? Is this a useful question to ask, and why?

4 Multiple regression

(Marks: 25)

In the same project as question 3, researchers had also asked whether the respondents had any children under 16. They created a single summary variable, 0 to mean no children, 1 for those who had children, and added that to the regression analysis. These were the new results:

- Constant: 35
- Coefficient for hours worked: -0.35
- Coefficient for having children: -7.5

1. Draw the two regression lines
2. Interpret the regression coefficients in words
3. Discuss the overall interpretation of this regression analysis, in comparison with the analysis without the children variable

5 Using Stata for Regression

(Marks: 20)

Using the same downloaded data set as in question 1, do the following separately for men and women:

- Create scatterplots of income (Y-axis) by hours (X-axis)
- Do you see association in these scatterplots?
- Get the correlation coefficient, and say what it means
- Run the regression analysis and write out the full $Y = a + b \times X$ equation
- Draw the line over the scatterplot, and interpret what it means

Having done this for both men and women, compare your results:

- What are the predicted earnings for men and women who work 40 hours?
- ... who work 20 hours?
- For which sex does the regression work better? Why might this be so?