*MSc in Migration Studies*
*Research Methods*
*Hilary Term 2015*

Reading (posted on WebLearn):
Teney, C., Jacobs, D., Rea, A. & Delwit, P. (2010): "Voting patterns among ethnic minorities in Brussels (Belgium) during the 2006 local elections", *Acta Politica*, 45 (3), 1-32

Stata commands
Several exercises require the use of Stata. For the commands for t-tests of means, two group tests of proportions and chi-square tests see pages 26 and 27 of the Stata guide. Please note that the command 'tabchi' is a user-written command and will need to be downloaded before it can first be used (see page 26).

Exercises
1. A researcher is interested in the effect of religiosity on quality of life. They suspect that religion increases people's life satisfaction, because it gives them a sense of purpose and the promise of an afterlife. They want to test this using the ESS. They operationalize religiosity by comparing people who aren't religious to those who are (question C17) and life satisfaction with B24 "*All things considered, how satisfied are you with your life as a whole nowadays? [..] where 0 means extremely dissatisfied and 10 means extremely satisfied.*"

1A. What is your dependent variable? What is its measurement level?

1B. What is you independent variable? What is its measurement level?

1C. What test should you use and why?

1D. What are the $H_0$ and $H_a$ for this test?

1E. It is a good habit to look at the data before running a statistical test: are there any unusual values in the dependent or independent variable? Is the relation between the two variables in the sample the same as you expected? For this latter step you can make a cross-tabulation (with row or column percentages) or calculate group-specific means.

1F. Run the test in Stata

1G. What is the p-value of the test statistic?

1H. What does this p-value signify?

1I. What do you conclude in relation to the hypotheses you set up? Use a significance level (alpha) of .05

1J. Does the data meet all the assumptions behind the test you performed?

2. In previous weeks we have looked at people's perception of discrimination in the justice system. A researcher would like to know if immigrants and natives differ in their judgements of the level of colour blindness of the justice system. They hypothesise that immigrants more often feel that the justice system is discriminatory than natives. They use ESS data, operationalising perception of the justice system using variable D30 and immigrant status on the basis of country of birth using C28.

   2A. What is your dependent variable? What is its measurement level?

   2B. What is your independent variable? What is its measurement level?

   2C. What test should you use and why?

   2D. What are the $H_0$ and $H_a$ for this test?

   2E. It is a good habit to look at the data before running a statistical test: are there any unusual values in the dependent or independent variable? Is the relation between the two variables in the sample the same as you expected?

   2F. Run the test in Stata

   2G. What is the p-value for the test statistic?

   2H. What does this p-value stand for?

   2I. What do you conclude in relation to the hypotheses you set up? Use a significance level (alpha) of .05

   2J. Does the data meet all the assumptions?

   2K. If applicable, do a residual inspection and interpret the results

   2L. Write up the test result in substantive terms; what can you conclude about the relation between immigrant status and perception of discrimination in the justice system?

3. Read the article by Teney et al (2010) on the political participation of minorities in Brussels.

   3A. What is the population of their study? And what is their sample? (see page 280-281)

   3B. What do the authors mean when they write that voters of certain parties were over- or underrepresented (p280)? What could have caused this over- or underrepresentation?

   3C. What do the authors mean when they write that "We cannot pool the data together for the three municipalities in an analysis of party choice" (p 281)? What does "pooling data" mean and why is it not possible here?

3D. They "collapse the results for smaller parties into a generic rest category together with void votes." (p 282). Why would they do that?

3E. What do the authors mean by "there is no statistically significant relationship between citizenship type (EU, non-EU or Belgian) and party preference"? (p 282)  In what kind of test (paired t-test of means, unpaired t-test of means, two-sample test of proportions, chi-square test) could they have determined this?

3F. Look at the 3 tables on page 283. What are the $H_0$ and $H_a$ for the chi-square tests?

3G. For table 1 explain what the p-value of the chi-square test statistic mean and draw a substantive conclusion about the relation between origin and party preference.

## Optional extra exercises

4.According to the New Economics of Labour Migration (NELM), it is relative not absolute deprivation that drives migration. A sense of relative deprivation can occur when people see how their relatives, friends or neighbours become richer after migration.  According to NELM, having migrants in one's network should lead to higher migration aspirations.  In the EUMAGINE survey we asked people about the presence of migrants in their network (questions mg1, mg10 and mg18) and about their international migration aspirations (question a1). Table 1 shows the results for Turkey. (the numbers are numbers of respondents).

A. What are the expected cell counts for the four empty cells if there is no relation between having (former) migrants in your network and international migration aspirations?

|  | No (former) migrants in network | Has (former) migrants in network | Total |
|---|---|---|---|
| Stay in this country |  |  | 1,199 |
| Go abroad |  |  | 801 |
| Total | 1,171 | 829 | 2,000 |

Table 1.  Migration aspirations by migrants in network. Source: EUMAGINE Turkey, unweighted data

B. What are the null-hypothesis and alternative hypothesis for a chi-square test of the relation between migration aspirations and having (former) migrants in one's network?

C. Table 2 displays the observed cell counts from the EUMAGINE survey. Calculate the chi-square and the degrees of freedom

|  | No (former) migrants in network | Has (former) migrants in network | Total |
|---|---|---|---|
| Stay in this country | 731 | 468 | 1,199 |
| Go abroad | 440 | 361 | 801 |
| Total | 1,171 | 829 | 2,000 |

Table 2.  Migration aspiration by migrants in network.  Source: EUMAGINE Turkey, unweighted data

D.  Look up the p-value with help from Stata. You can do this using the following command
    *. dis chi2tail(df, value of chisquare)*
    So for example, if you have determined that the df=3 and the value chi-square you calculated in 4C
    is 5.9, you type
    *. dis chi2tail(3, 5.9)*
    The value Stata returns is that of the p-value (probability) of drawing a sample in which the
    distribution is at least as disproportionate as that presented in Table 2, if there is no relation
    between network and migration aspiration in the population (i.e. 18-39 years olds in four regions of
    Turkey).

E.  What do you conclude at a significance level of alpha=.05?


5.  Immigration reform is a contentious political issue in the US. It is often argued that though the rapidly
    increasing Hispanic population in the US is social conservative and would therefore be likely to be
    politically aligned with the Republican party, the Republican's restrictive stance on immigration deters
    Hispanic voters. The hypothesis is that immigrants are more likely to be in favour of open migration
    policies, because that would allow their friends and family members to migrate. A researcher would
    like to explore the attitudes in Europe. He wants to test whether immigrants in Europa have less
    restrictive attitudes towards immigration than natives and decides to test this with ESS data. Because
    he is mostly interested in the share of people who would like to stop immigration he operationalises
    "attitude towards immigration" by way of the dichotomous variable you created for question 3 of the
    week 5 exercises *(Using variable B36, make a new variable (immstop) that is '1' for respondents who
    want no more immigration, and '0' for respondents who want to allow many, some or a few.)*, and
    immigrant status using variable C28 (just like in the exercises above).

    5A.  What is your dependent variable? What is its measurement level (nominal, ordinal,
         interval/ratio)?

    5B.  What is your independent variable? What is its measurement level?

    5C.  What test should you use and why?

    5D.  What are the $H_0$ and $H_a$ for this test?

    5E.  Run the test in Stata

    5F.   Look at the output from Stata. What is the p-value for the test statistic? (so for the z-score, t-score
          or chi-square score)

    5G.  Write up the interpretation of the p-value in statistical terms

    5H.  If applicable do a residual inspection.

    5I.   Does the sample meet all the assumptions of the test?

    5J.   Write up the test result in substantive terms; what can you conclude about the difference in
          attitudes towards restrictive migration policies of immigrants and natives?