

**Human Sciences A**  
**Statistics Exercises 2017-18**

**Relevant Units**

Block 3 Units on Quantitative Analysis (recommended units are presented for each exercise)

**Additional Textbooks:**

Pallant, J. (2013). *SPSS survival manual (5<sup>th</sup> Ed.)*. Maidenhead, UK: Open University Press (This book contains a basic introduction to SPSS, and there are good instructions on how to use SPSS)

Field, A. (2013). *Discovering statistics using SPSS (4<sup>th</sup> Ed.)*. London: Sage Publications Ltd (This is slightly more advanced, good companion website)

**Instructions:**

- An SPSS file containing data from a survey has been uploaded on Loop; please download this file.
- There are **3 statistics exercises** in total to be completed, using SPSS to analyse the data
  - Exercise 1: Describing and Summarising Data (20%)
  - Exercise 2: Tests of Difference (30%)
  - Exercise 3: Analysing categorical data and Multivariate Analysis (50%)
- Use the forum on Loop to ask questions and discuss the exercises with other students.
- While you may ask for help and work with other students you should ensure that you complete the full set of exercises on your own. Do not under any circumstances copy output from other students!
- **Submission:**
  - Compile the three exercises into **one Microsoft Word** document with an additional document containing relevant SPSS output, to be submitted by **Monday 4<sup>th</sup> December 2017**.
- **Layout document:**
  - Your statistics exercises document should be laid out in **three sections** relating to each of the three exercises. Each question number and its related answer should be *clearly labelled*.  
Label the file as **'YournameStatsExercises.doc'** [e.g., **JohnSmith\_StatsExercises.doc**].
  - Compile SPSS output for each exercise into a separate document. Please note that the output should only contain the tests relevant to the exercise in question (i.e., you should delete any test conducted by mistake etc.). Label the Output File as **'JohnSmith\_StatsOutput.spv'**

## BACKGROUND TO DATASET

### Questionnaire Survey

The data used for these exercises were collected as part of a recent survey of student **attitudes to research** in three institutions. The students were informed that some of the data they provided would be used by other students to carry out statistical exercises.

The questionnaire comprised five sections:

- ☐ **Section 1:** Collects information on experience of research; respondents were asked to indicate, yes, no, or don't know.
- ☐ **Section 2:** Contains 25 statements about research. Respondents were asked to indicate their level of agreement or disagreement with the statements.
- ☐ **Section 3:** Contains 10 statements about mathematics. Respondents were asked to indicate their level of agreement or disagreement with the statements.
- ☐ **Section 4:** Asks for personal details including gender, age, Leaving Certificate results, institution and intentions with regard to postgraduate study.
- ☐ **Section 5:** Open ended questions (these have been removed from the dataset)

### Preparing Data for Analysis

The data have been coded and inputted into an SPSS data file which is on Loop.

### Variables in Questionnaire

The variables in the questionnaire are presented in Appendix A at the end of this document.

### Exercise 1:

#### Describing and Summarising Data (20%)

**Relevant Reading:** Block 3 Unit 2 and chapter from textbook

Part 1: Choose the correct answer for each of the following:	
1.	A null hypothesis ... a) States that the experimental manipulation will have an effect b) Is rarely used in experiments c) Predicts that the experimental manipulation will have no effect d) None of the above
2.	In conducting a survey, you gather data on age. Participants are asked to select from: 18-25; 25-40; 41-65; 66 and over. Is this variable ...? a) Scale b) Ordinal c) Nominal d) None of the above
3.	A Type I error is when: a) We conclude that there is a meaningful effect in the population when in fact there is not b) We conclude that there is not a meaningful effect in the population when in fact there is c) We conclude that the test statistic is significant when in fact it is not d) The data we have entered into SPSS is different to that collected
4.	What does a positive value of skewness indicate? a) The distribution is pointed with many scores in the tails b) The distribution is flat with few scores in the tails c) The scores are piled up on the left hand side of the distribution d) The scores are piled up on the right hand side of the distribution
Part 2: Answer the following in relation to distribution of data:	
5.	Briefly explain the following: a) What is meant by measures of central tendency b) What is meant by mean, mode and median

6.	Calculate. and write a short interpretation of, the mean, median and mode for the variable <i>Age</i> in the dataset provided
<b>Part 3: Using SPSS:</b>	
7.	Produce a histogram, showing the normal curve, for the variable <i>CAO Points</i>
8.	Calculate the skewness and kurtosis values for the variable <i>CAO Points</i>
9.	Briefly discuss the distribution of the <i>CAO Points</i> variable in the sample in terms of measures of central tendency and dispersion. Finally, suggest ways in which the data could be recoded to make it amenable for statistical testing.
<b>Part 4: Using SPSS calculate and provide a summary for the following:</b>	
10.	<p>Percentage of students attending each University</p> <p><b>Note:</b> Your interpretation should include both a table and a short verbal interpretation.</p>
11.	<p>Percentage of students who engaged in different types of research activities listed in questions Q1.1 - Q1.13</p> <p><b>Note:</b> This task requires you to produce a table of frequencies. The 13 items should be listed in one table, together with a brief commentary of what the frequencies tell us about the students' experience of research.</p>

## Exercise 2: Tests of Difference (30%)

**Relevant Reading:** Block 3 Units 3, 4, and 5, and chapters from textbook

*Hints:*

- *You should also organise your data before starting to answer these questions.*
- *Use Explore in SPSS to illustrate your answer with statistics, histograms and boxplots.*

Part 1: Choose the correct answer for each of the following:	
1.	Which of the following is an independent subjects design? a) All participants perform in each condition b) Each participant is tested twice, once in each condition c) Different participants perform in each condition d) None of the above
2.	You are interested in the effects of teaching methods on student learning. A group of students are randomly assigned to receive tuition via large lectures, small tutorial groups or online podcasts. At the end of the term, all students take an exam to assess their knowledge of the subject matter. How many levels does the independent variable have? a) One b) Two c) Three d) Four
3.	An advantage of repeated measures designs is: a) Absence of practice effects b) Greater control over potential confounding variables c) Greater variance in outcomes d) Less time is required to gather data
Part 2: Tests of Difference – Independent Subjects Design	
4.	<p>The researchers who collected the data available in your SPSS file were interested in looking at potential gender-related differences. Thus, the following research question (RQ) was formulated:</p> <p style="text-align: center;">Is there a difference between males and females in terms of research anxiety scores (<b>NOT</b> ResearchPost)?</p> <p>For this research question, you will need to create a new variable, <i>Research Anxiety Score</i>. Instructions are uploaded on Loop which explain how to compute</p>

	<p>this variable, which you need to do to proceed.</p> <ol style="list-style-type: none"> <li>What are the independent and dependent variables? How many levels are in each of these variables?</li> <li>State the alternate and null hypothesis for the research question.</li> <li>What kind of data is gathered for the dependent variable?</li> <li>Which type of test is appropriate – parametric or non-parametric? Briefly explain your reasoning</li> </ol>
5.	<p>Using SPSS, conduct the relevant test to address the research question. Analyse the output briefly, addressing the following:</p> <ol style="list-style-type: none"> <li>Significance and report the test value (where appropriate)</li> <li>The implication of this result in relation to the variables under investigation</li> <li>How confident can we be that this result is accurate (i.e., not due to Type 1 error) (Hint: look at the <math>CI_{95}</math> where appropriate).</li> </ol>
<p style="text-align: center;"><b>Part 3: Tests of Difference – Repeated Measures Design</b></p>	
6.	<p>The previous RQ compared the research anxiety scores between males and females. Now we want to look at the impact of an experimental intervention on research anxiety scores, involving additional staff consultation hours. At the end of the intervention, respondents were asked to fill in the research anxiety questionnaire again. Their scores for the post Intervention Research Anxiety Scale were calculated, and these were recorded in the SPSS file as the variable <b>ResearchPost</b>.</p> <p>In the following question, the researchers were interested in looking at differences across time and gender differences. Thus, the following research questions were posed?</p> <p>Is there a difference between males' research anxiety scores between Time 1 and Time 2? Is there a difference in females' research anxiety scores Time 1 and Time 2?</p> <ol style="list-style-type: none"> <li>What are the independent and dependent variables? How many levels are in each of these variables?</li> <li>State the alternate and null hypothesis for the research question.</li> <li>What kind of data is gathered for the dependent variable?</li> <li>Which type of test is appropriate – parametric or non-parametric? Briefly explain your reasoning</li> </ol>
7.	<p>Using SPSS, conduct the relevant test to address the research question. Analyse the output briefly, addressing the following:</p> <ol style="list-style-type: none"> <li>Significance and report the test value (where appropriate)</li> <li>The implication of this result in relation to the variables under investigation</li> <li>How confident can we be that this result is accurate (i.e., not due to Type 1 error) (Hint: look at the <math>CI_{95}</math> where appropriate).</li> </ol>

### Exercise 3: Analysing categorical data and Multivariate Analysis (50%)

**Relevant Reading:** Block 3 Units 6 and 9 and Chapters from Textbook.

<b>Part 1: Choose the correct answer for each of the following:</b>	
1.	What is the lowest expected frequency in any cell for a chi-square test for independence? a) 1 b) 2 c) 5 d) 10
2.	How many dependent variables does a two-way ANOVA have? a) One b) Two c) Three d) Four
3.	If a study has employed a two-way mixed ANOVA, what does this mean? a) There were male and female participants in the study b) There were two factors in the study, each with a different number of levels c) There was a repeated measures factor and an independent measures factor in the study
4.	Chi-square test for independence allows the researcher to: a) Test for differences between two categorical variables; b) Predict the effect of one variable upon another; c) Test for differences between two continuous variables;
<b>Part 2: Analysing categorical data</b>	
5.	<p><b>Note:</b> Read through the unit and relevant reading first, then organise your data for this analysis. You will recode the variable <i>Grademaths</i> into a new variable: <i>Mathslevel</i>. Instructions on how to do this have been uploaded onto Loop.</p> <p>Is there an association between male and female students and the level of mathematics taken at leaving certificate?</p> <p>a) State the alternate and null hypothesis b) What is the appropriate test statistic to report if the test is significant?</p>

6.	<p>Conduct the chi-square test using SPSS. Analyse the output and report the following findings:</p> <ul style="list-style-type: none"> <li>a) Percentage of males and females in each category; edit the contingency table to show row and column totals, as well as % within gender</li> <li>b) Report the significance and test value</li> <li>c) If the result is significant, report the effect size</li> <li>d) What conclusions do you draw from the results?</li> <li>e) If the distribution of frequencies does not permit you to do a chi-square test, suggest ways of recoding the variable to make it amenable to this test.</li> </ul>
<p style="text-align: center;"><b>Part 3: Comparing Multiple Groups – Analysis of Variance</b></p>	
7.	<p><b>Note:</b> Read through the unit and relevant reading first, then organise your data for this analysis. You will need to recode the age variable into groups using SPSS. Instructions on how to do this have been uploaded onto Loop.</p> <p>In the following exercise, the researchers were interested in looking at differences in anxiety scores across age groups. Thus, the following research question was posed:</p> <p style="padding-left: 40px;">Is there a difference in research anxiety scores for students who are under 30yrs, 30-39yrs and 40+yrs?</p> <ul style="list-style-type: none"> <li>a) What are the independent and dependent variables? How many levels are in each of these variables?</li> <li>b) State the alternative and null hypothesis</li> <li>c) What is the appropriate statistical test to perform to address this RQ?</li> <li>d) What are the main assumptions underlying this test? Briefly explain how these assumptions are met in the current data set.</li> </ul>
8.	<p>Conduct the appropriate test using SPSS. Analyse the output under the following headings:</p> <ul style="list-style-type: none"> <li>a) Significance and report the test value (where appropriate)</li> <li>b) The implication of this result in relation to the variables under investigation</li> <li>c) iii) How confident we can be that this result is accurate (i.e. not due to Type 1 error).</li> </ul>



## Appendix A

### Codebook for Questionnaire Variables

#### Part 1: Experience of Research

		Yes	No	Don't Know
Q1.1	Heard lecturer discuss their research	1	2	3
Q1.2	Heard guest lecturer discuss their research	1	2	3
Q1.3	Read research paper or report written by lecturer	1	2	3
Q1.4	Attended research seminar university/college	1	2	3
Q1.5	Attended a research conference	1	2	3
Q1.6	Participated in a research project run by lecturer	1	2	3
Q1.7	Courses in research skills	1	2	3
Q1.8	Independent project as a part or whole of a module	1	2	3
Q1.9	Dissertation or thesis	1	2	3
Q1.10	Took part in research related fieldwork activities	1	2	3
Q1.11	Acted as a research assistant	1	2	3
Q1.12	Contributed to research conference paper or poster	1	2	3
Q1.13	Contributed to a research paper/report	1	2	3

## Part 2: Attitudes to Research

		Strongly agree	Agree	Neutral	Disagree	Strongly Disagree
Q2.1	Research skills will be useful for my future career	1	2	3	4	5
Q2.2	Experiencing research is an important part of the university experience	1	2	3	4	5
Q2.3	My lecturers integrate research findings into their teaching very well	1	2	3	4	5
Q2.4	Research skills are essential to my field of study	1	2	3	4	5
Q2.5	Research skills should be taught to all students	1	2	3	4	5
Q2.6	I would like to learn more about research methods	1	2	3	4	5
Q2.7	Research skills are useful for every professional	1	2	3	4	5
Q2.8	I would like to use research skills in my career	1	2	3	4	5
Q2.9	I am anxious about using research methods such as surveys and interviews	1	2	3	4	5
Q2.10	The thought of carrying out research on my own makes me apprehensive	1	2	3	4	5
Q2.11	I feel anxious about using statistics to analyse data	1	2	3	4	5
Q2.12	The thought of producing an independent research project is stressful	1	2	3	4	5
Q2.13	I feel I have learnt a lot about research in my studies	1	2	3	4	5
Q2.14	I think research is very time consuming	1	2	3	4	5
Q2.15	I think that carrying out research is difficult	1	2	3	4	5
Q2.16	I would prefer to do examinations than research projects	1	2	3	4	5
Q2.17	I would like to develop field research skills (e.g. observations, measurements etc)	1	2	3	4	5
Q2.18	Research is boring	1	2	3	4	5

		Strongly agree	Agree	Neutral	Disagree	Strongly Disagree
Q2.19	I enjoy carrying out research	1	2	3	4	5
Q2.20	I would like to carry out research using surveys and interviews	1	2	3	4	5
Q2.21	I enjoy reading about new research	1	2	3	4	5
Q2.22	Most students benefit from learning research methods	1	2	3	4	5
Q2.23	Research-orientated thinking plays an important role in everyday life	1	2	3	4	5
Q2.24	Research-oriented thinking does not apply to my personal life	1	2	3	4	5
Q2.25	I am confident that I can master research skills	1	2	3	4	5

### Part 3: Experience of Mathematics

		Strongly agree	Agree	Neutral	Disagree	Strongly Disagree
Q3.1	I hated mathematics in school	1	2	3	4	5
Q3.2	I was good at mathematics in school	1	2	3	4	5
Q3.3	My mathematics teacher helped me to understand mathematics	1	2	3	4	5
Q3.4	My parents were able to help me with mathematics	1	2	3	4	5
Q3.5	My school friends thought mathematics was boring	1	2	3	4	5
Q3.6	My school friends thought mathematics was difficult	1	2	3	4	5
Q3.7	I enjoyed working with my school friends on mathematics problems	1	2	3	4	5
Q3.8	I am confident that I can master statistics	1	2	3	4	5

**Part 4: Personal Details**

Gender	1 = Male; 2 = Female
Age	Actual Age
Institution	1 = University A; 2 = University B; 3 = University C
Second Level qualification	1 = Incomplete 2nd Level; 2 = Leaving Cert; 3 = Leaving Cert equivalent; 4 = Other;
Grade in Leaving Cert	Maths 1 = A Hons; 2 = B Hons; 3 = C Hons; 4 = D Hons; 5 = A Ord; 6 = B Ord; 7 = C Ord; 8 = D Ord; 9 = Other
Number of CAO points	Actual number of points
Intentions for postgrad study	1 = Planning to do Postgrad; 2 = Not planning to do Postgrad