D. Clin. Psy. 2\textsuperscript{nd} Year Statistics Examination 2011

Maximum time allowed \textbf{3 hours}

\textbf{Instructions}
Please read the following instructions carefully before beginning the exam:

1. You are not allowed to communicate in any way with other candidates during the exam.
2. You are allowed to use textbooks and lecture notes if you wish.
3. Please do not change the dataset in any way unless explicitly told to do so. Although you may be asked to identify outliers or suggest transformations, \textit{please do not actually make any such changes unless explicitly told to do so.}
4. Please write your candidate number clearly at the top of each page of your answer booklet.
5. You must answer all questions. Failure to answer a question will mean a mark of zero being given for that question. Always attempt to answer all questions.
6. The numbers in brackets at the end of each question refer to the basic marks available for that question. Bear this in mind when answering. Please note however that there will not necessarily be a one-to-one correspondence between the number of marks available and the apparent number of specific answers to any one question. For example, harder questions may be assigned more marks than easier ones, regardless of the number of responses required.
7. Some questions will include brackets that indicate a possible plural, e.g. statistic(s). This in no way indicates that a plural is or is not expected.
8. If you have any questions about the exam, or you need more paper, quietly raise your hand and wait for an invigilator to come to your desk.
9. Please write clearly on the answer sheet.
10. The amount of space given for a question is not an indication of the length of answer expected.
11. All questions should be marked on the answer sheet. If you do not have enough space to finish your answer please continue in the \textit{additional paper} booklet, marking the question number carefully. If you use extra paper please ensure that each page is marked with your candidate number and that all sheets are tied together with the answer booklet.
12. Please do not write answers purely in list form or in shorthand, unless you are forced to do so due to constraints of time. Marks will be given for clear and concise answers.
13. When you have finished answering all the questions please raise your hand and wait for an invigilator to check your answer sheet. You may then leave.
14. Questions 1 – 10 do not require SPSS. Only question 11 requires you to use SPSS.
15. Note that 52\% of the possible marks are allocated to question 11.
Please answer all the questions in this exam. Remember that your answers need only be brief. This exam is a test of your knowledge of statistics and your practical knowledge of SPSS not a test of critical skills.

Please also note that there are a maximum of 56 marks to be gained in this exam. Therefore, as a rough guide, you should aim to spend no more than 3 minutes on a question worth 1 mark.

Section I
Note: This section does not require you to use SPSS

Multiple Choice Questions
Please circle the correct answer below.

1. As sample size increases, power will tend to decrease.
   True False [1]

2. In a positively-skewed distribution, scores above the mean will be more spread out than scores below the mean.
   True False [1]

3. Levene's test for homogeneity of variances can be used to assess whether a sample is normally distributed.
   True False [1]

4. Logistic regression is used to investigate the predictors of binary outcomes.
   True False [1]

5. Only correlations greater than 0.3 can be considered statistically significant.
   True False [1]
6. A colleague has been asked to redo an analysis, correcting for multiple comparisons due to concerns about inflation of Type I error. They do not understand what this means. Please describe how you would explain the concepts to them, with examples. [5]
7. Imagine you are working in a forensic CAMHS service. On your waiting list 15% of clients are current drug users and 18% have in the past been arrested by the police. Assuming these factors are independent, what is the probability that your next randomly chosen patient will be a drug user who has never been arrested? [3]

8. Some parametric tests are subject to an assumption of normality. Why is this, and what are the consequences of violating them? Discuss with reference to the central limit theorem, and provide an example. [4]
9. A study was conducted to investigate whether problem-solving skills and beliefs could predict worrying in primary school children. 806 children (8-11y) were screened using the Penn State Worry Questionnaire for Children (PSWQ-C). The children then completed the Problem-Solving Inventory (PSI), which provides three subscale scores: (i) problem-solving confidence, (ii) approach-avoidance style, and (iii) personal control (all three are continuous measures). Age and gender were also recorded. How would you analyse this study? Please state the name of the statistical technique(s) and the independent and dependent variables. Please also describe the kinds of conclusions the analysis would allow the researcher to come to. Can you think of any weaknesses in the study's design?
10. In a study of autobiographical recall in borderline personality disorder (BPD), three groups were recruited (i) depressed (N=22) and (ii) non-depressed (N=9) patients who met criteria for BPD were matched by age and gender with (iii) healthy controls (N=29). All were assessed with the Autobiographical Memory Test (AMT), which produces a single continuous score. The National Adult Reading Test (NART, also a continuous score) was also administered to allow the investigators to control for IQ in their analysis. How would you analyse this study? Please state the name of the statistical technique(s) and the independent and dependent variables. Please also describe the kinds of conclusions the analysis would allow the researcher to come to. Can you think of any weaknesses in the study's design? [4]
Section II

11. Introduction: This section requires that you use SPSS for Windows. Read each question carefully and write your answers in the spaces provided. As before, if you need to continue beyond the space available please use the additional paper provided, marking the question number clearly in the margin.

When conducting statistical tests for this exam, it is not necessary to consider issues of normality and other assumptions relating to the specific test, unless explicitly instructed to do so.

The data file you have been given (exam11.sav) contains data from a study conducted to investigate whether cognitive stimulation might improve symptoms in people with early-diagnosis Alzheimer’s who were also suffering from depression. At intake to the study, age and educational attainment (education: 0=ceased education before 18, 1=attained a post-18 qualification) were recorded along with baseline measures of depressive symptoms (BDI1) and verbal learning (VLT1). The patients then underwent a two-month programme of weekly museum visits and assessed reading material, after which the measures were repeated (BDI2 and VLT2). Difference scores have been calculated for depressive symptoms (BDI_d) and learning (VLT_d). The number of self-initiated outpatient hospital visits during the period of the study was also noted (HospVisits: 0=none, 1=one, 2=more than one). All measures can be treated as continuous variables, other than education and HospVisits which are categorical.

a) Consider the variable representing age (age). Make brief notes about the distribution of the variable in relation to issues of normality and outliers. Would you suggest any remedial changes to the variable, and if so, what would they be (do not apply them to the data in SPSS)? [3]
b) Is there a relationship between age (age) and baseline verbal learning (VLT1)? Please quote the name of the statistical test you did, the relevant statistic(s) and the probability values(s). If the test is significant, comment on the direction/source of the effect. [2]

c) Consider the variable representing baseline verbal learning (VLT1). Working separately for the educational attainment groups (ie using the variable education to select cases), make brief notes about the distribution of the variable in relation to issues of normality and outliers. Would you suggest any remedial changes to the variable, and if so, what would they be (do not apply them to the data in SPSS; if you have selected cases, remember to unselect afterwards)? [4]
d) Is there a relationship between baseline depression symptoms (BDI1) and baseline verbal learning (VLT1)? Please quote the name of the statistical test you did, the relevant statistic(s) and the probability values(s). If the test is significant, comment on the direction/source of the effect. [2]
e) Is the change in depression symptoms (BDI_d) different for the two education attainment groups? Please quote the name of the statistical test you did, the relevant statistic(s) and the probability value(s). If a test is significant, comment on the direction/source of the effect. [3]

f) Conduct a statistical test to determine whether there is significant change in learning performance between the intake (VLT1) and the end of the intervention (VLT2). If the test is significant, comment on the direction and meaning of the effect. [3]
g) Perform a statistical analysis to see whether learning performance changing over time (VLT1, VLT2) interacts with education. Include a plot to illustrate the findings. Please quote the name of the statistical test you did, the relevant statistic(s) and the probability values(s). If a test is significant, comment on the direction/source of the effect. [4]
h) Carry out a statistical test to see whether there are differences in learning improvement (VLT_d) in the different hospital admission groups (HospVisits). Please quote the name of the statistical test that you used to do this and the relevant statistic(s) and probability value(s). Please comment on the meaning of these results. [2]

i) Are age and change in depression (BDI_d) predictive of improvement in learning (VLT_d)? Quote any relevant statistics and describe the test you used. Now repeat this analysis adding education as a predictor. Do the results change? [3]
j) Please discuss the overall pattern of findings in this series of analyses. What conclusions can be drawn about the efficacy of cognitive stimulation in Alzheimer’s disease? What are the weaknesses in the study, and how might these be addressed? [3]
End of examination

Please check your answers carefully. Please also make sure that your candidate number is written clearly at the top of every page of your answer booklet and if you have used any extra paper please make sure that every page is labeled with your candidate number and that these extra sheets are tied together with your answer booklet.