University of Toronto Scarborough
Department of Computer and Mathematical Sciences
STAC32 (K. Butler), Midterm Exam
October 19, 2019

Aids allowed (printed or handwritten): My lecture overheads (slides); Any notes that you have taken in this course; Your marked assignments; My assignment solution; Non-programmable, non-communicating calculator.

This exam has 10 numbered pages of questions. Check to see that you have all the pages. There is an additional empty page that you can use if you need more space for any answers.

In addition, you should have an additional booklet of output to refer to during the exam. Contact an invigilator if you do not have this.

Answer each question in the space provided (under the question).

The maximum marks available for each part of each question are shown next to the question part.

You may assume throughout this exam that the code shown in Figure 1 of the booklet of code and output has already been run.

The University of Toronto's Code of Behaviour on Academic Matters applies to all University of Toronto Scarborough students. The Code prohibits all forms of academic dishonesty including, but not limited to, cheating, plagiarism, and the use of unauthorized aids. Students violating the Code may be subject to penalties up to and including suspension or expulsion from the University.

Question 1 (11 marks)

When you drive along the highway, you might notice a lot of vegetation (grass, plants etc.) along the side of the highway, in the middle of on-ramps, etc. This vegetation has to be attended to: in particular, if it gets more than 30 cm tall, it is dangerous because car drivers cannot see what is on the other side. Highway authorities need to devise a system to keep the vegetation less than 30 cm tall without spending a lot of money on maintenance. In Texas, an experiment was run. The vegetation next to a typical highway was divided into 36 sections. The equipment used to mow the vegetation has an adjustable blade which can be adjusted to a height of 5, 10, and 20 centimetres (0.05, 0.10, and 0.20 metres). In addition, the vegetation can be mowed 1, 2, or 3 times a year. Each section of the highway was randomly assigned one of these blade heights and mowing frequencies (in such a way that each combination was used on four sections of the highway). One year after the vegetation was last mowed, the maintenance crew came back and measured the maximum height of vegetation on that section of the highway.

The data file is shown in Figure 2 (in the Booklet of Code and Output), and this data file is also stored in `mowing.txt` in your current project in R Studio Cloud (or, if you prefer, in the folder of your current project in R Studio on your computer).

(a) (3 marks) Using something from the `tidyverse`, give R code to read this data into a data frame called `mowing`.

(b) (1 mark) What R code would display (at least some of) the values in your data frame?

(c) (2 marks) The column `height` is really a number, but ought to be treated as a categorical variable. Why is that? Explain briefly.

(d) (2 marks) Give R code to make an appropriate plot of `height` and `vegetation`, ignoring (for this part) `frequency`.

(e) (3 marks) We now want to make a suitable graph that includes all three variables. Describe the kind of graph that you would draw, and give R code to produce it.

Question 2 (13 marks)

People are concerned about the use of nitrates as meat preservatives. One study looked at possible effects of these chemicals. Bacteria cultures were grown in a medium containing nitrates, and the rate of uptake of radio-labelled amino acids was then determined for each culture. The data are shown in Figure 3, in units of disintegrations per minute. The data are shown in Figure 3.

It is known that the mean rate of uptake for cultures grown without nitrates is 8000 in these units. We will be investigating whether the addition of nitrates results in a decrease in mean uptake rates.

(a) (2 marks) A histogram is shown in Figure 4. Give the code that was used to produce this histogram. (The data frame is called `nitrates`.)

(b) (2 marks) The researchers wanted to assess the effect of nitrates on "typical" uptake rate. They ran a t-test rather than a sign test. Why do you think they decided to do this? Explain briefly.

(c) (3 marks) Give code to obtain a suitable t-test.

(d) (2 marks) The output from your t-test is shown in Figure 5. What do you conclude from it, in the context of the data?

(e) (2 marks) What *two* changes to the code for your hypothesis test would produce output containing a 95% confidence interval for the true mean uptake rate? (If you only think you need one change, explain briefly why your one change is sufficient.)

(f) (2 marks) A 95% confidence interval for the population mean uptake rate is shown in Figure 6. The researchers thought that maybe adding nitrates to the bacterial culture might reduce the mean uptake rate by 500. By looking at the confidence interval, do you think that the researchers would have had a reasonable chance of being able to prove that the mean uptake rate was less than 8000, using the sample size that they had? Explain briefly.

Question 3 (14 marks)

In a power plant, water is used for cooling, and the water is then discharged into a nearby river. It has been determined that as long as the mean temperature of the discharged water is no more than 65 degrees Celsius, there will be no negative effects on the river's ecosystem. To find out whether the power plant is discharging water that is too warm, a scientist will take 50 water specimens (at randomly selected times) and record the temperature of each one. If μ denotes the mean temperature of all the discharged water, the scientist will then test $H_0 : \mu = 65$ against $H_a : \mu > 65$.

(a) (2 marks) Describe a type I error in this context.

(b) (2 marks) Describe a type II error in this context.

(c) (4 marks) The population standard deviation of water temperature measurements is believed to be about 15 degrees. Give R code to **estimate** the power of the test when the mean water temperature is actually 68 degrees. Assume that water temperature readings have at least approximately a normal distribution.

(d) (2 marks) The output from your code is shown in Figure 7. What is your estimated power for this test?

(e) (2 marks) The scientist's manager would prefer to design the test to have a power of 0.7. Using Figure 8, approximately how many water specimens would the scientist need to take to achieve this?

(f) (2 marks) Explain briefly why your answer to the previous part compares with the original number of water specimens in the way that it does.

Question 4 (8 marks)

In each of the situations below, say whether you would use a *one-sided* or *two-sided* test, and also whether you would use a *one-sample* or *two-sample* test. In each case, therefore, you need to say something about both the number of sides and the number of samples, and give a brief justification of your choices.

(a) (2 marks) A study of wait times in coffee shops was carried out in Boston. The researchers were concerned about whether females had to wait longer than males on average. They observed a number of customers, and for each one recorded the wait time in seconds and whether each customer was male or female.

(b) (2 marks) A machine part has a hole in it that is supposed to be exactly 5 centimetres in diameter. The machine part is produced by a process that can be adjusted to produce parts with holes of different sizes. The process supervisor takes a sample of 10 parts (from the process at its current settings) and finds the mean and standard deviation of hole sizes of these parts. If the mean hole size is too far away from 5 cm, the process supervisor will need to adjust the process settings.

(c) (2 marks) The instructor of a large introductory psychology class believes that students need to spend 10 hours studying for the final exam to master the material, and is concerned that students are studying less than they should. The student newspaper reported that, for a random sample of 411 students in the course, the mean time spent studying for the final exam was 7.74 hours with a standard deviation of 3.40 hours.

(d) (2 marks) Forty men were recruited from a dating site. Each man has a profile in which he reports his height. The researchers also recorded each man's actual height, and compared it with the height reported in that man's profile. The researchers were trying to find out whether men on this dating site systematically reported themselves as being taller than they actually were.

(If you think this one will be neither a one-sample nor a two-sample test, describe what you think it is instead.)

Question 5 (7 marks)

Medical research has shown that repeated wrist extension beyond 24 degrees increases the risk of hand and wrist injuries. Some students at Cornell University were given a proposed new mouse design. While using the mouse, each student's wrist extension was measured. Our interest is in whether the average wrist extension is greater than 24 degrees, where the average could be the mean or median.

(a) (4 marks) A histogram of the wrist extension values is shown in Figure 9. Two possible analyses of these data are shown in Figures 10 and 11. Which of these analyses do you prefer and why, and what do you therefore conclude in the context of the data?

(b) (3 marks) If you don't remember `filter`, look at Figure 12 to see how it works.

Some further analysis is shown in Figure 13. By comparing this Figure with Figures 10 and 11, what do you learn about the behaviour of the t-test and sign test, and what general principle does it illustrate? Explain briefly.

Question 6 (11 marks)

Is it true that learning to play chess can improve your memory? In a study, sixth-grade students who had not previously played chess took weekly chess lessons and played chess daily for 9 months. Each student took a memory test called the "Test of Cognitive Skills" before starting the chess program and again at the end. The data are shown in Figure 14, with `pre_test` and `post_test` denoting the scores for each student before and after the chess program (respectively). The data frame is called `chess`.

(a) (2 marks) These are matched-pair data. How can you tell? Explain briefly.

(b) (3 marks) Give code to run a suitable t-test on these data.

(c) (3 marks) What code would produce a suitable graph for assessing whether the t-test you just did was appropriate?

(d) (3 marks) Why might you have guessed, looking at the data in Figure 14, that a suitable sign test for comparing memory test scores from before and after the program would have produced a significant result? Explain briefly.

Question 7 (10 marks)

Some people have the ability to remember accurately vast amounts of information about themselves, without using mnemonic tricks or extra practice. This ability is called "Highly Superior Autobiographical Memory" or HSAM. A study recruited adults with diagnosed HSAM and also control individuals of similar age without HSAM. The aim of the study was to determine what makes HSAM work. All the subjects in the study were given a large number of cognitive and behavioral tests. Some of the results for a visual memory test are shown in Figure 15. A higher score is better.

(a) (3 marks) What code would run a suitable Mood's median test on these data? (You may assume `library(smmr)` has already been run.)

(b) (3 marks) The output from your Mood's median test is shown in Figure 17. What do you conclude from this output, in the context of the data?

(c) (1 mark) There are 29 observations in the data set, but only 23 in the table in the Mood's median test in Figure 17. What happened to the others?

(d) (3 marks) A pair of graphs is shown in Figure 18. Explain briefly what you conclude from this, and thus discuss whether we should have run a Mood's median test or whether some other test would have been better instead. If you think some other test would have been better, give the name of the test you would prefer. (If you would prefer to see some other plot to help you decide, describe what you would like to see and why.)

Question 8 (11 marks)

Farmers know that driving heavy equipment like tractors over the soil, especially if the soil is wet, compresses the soil and makes it more difficult for crops planted in that soil in the future to grow. One way of quantifying this is to measure something called "penetrability", which is a measure of how much resistance plant roots will meet when they try to grow through the soil. On the scale measured, a high penetrability means that plants find it easier to grow.

A study was carried out at a research station. An area of soil was divided into three plots A, B, C. (These are plots of land, not `ggplot` plots). Plot A was driven over by a tractor in wet weather. Plot B was driven over by a tractor in dry weather. Plot C was left as it was. 20 locations were chosen at random within each plot and the soil penetrability measured. Some randomly chosen rows of the data are shown in Figure 19.

(a) (3 marks) Some plots are shown in Figures 20 and 21. Use either or both of these Figures to assess the two major assumptions for analysis of variance (that `aov` would run). In your explanation, make sure to mention which plot you are drawing each conclusion from.

(b) (2 marks) Figures 22, 23 and 24 show three possible analyses of these data. Which one of these analyses do you think is the most appropriate? Explain briefly. Your answer should contain a Figure number.

(c) (4 marks) What do you conclude from your chosen analysis, in the context of the data? Note that each analysis has a Part (i) and a Part (ii). Your explanation should include a discussion of what you conclude from Part (i), whether or not you need to do Part (ii), and (if appropriate) what you conclude from Part (ii).

(d) (2 marks) Do you think your conclusions would make sense to farmers? Explain briefly.

Use this page if you need more space to write your answers. Be sure to label any answers here with the question and part that they belong to.