University of Toronto Scarborough
Department of Computer and Mathematical Sciences
STAD29 (K. Butler), Midterm Exam
March 2, 2024

Aids allowed (on paper, no computers):

- My lecture overheads (slides)
- Any notes that you have taken in this course
- Your marked assignments
- My assignment solutions
- Non-programmable, non-communicating calculator

This exam has 7 numbered pages of questions plus this cover page.
In addition, you have an additional booklet of Figures to refer to during the exam.
The maximum marks available for each part of each question are shown next to the question part.
If you need more space, use the last page of the exam. Anything written on the back of the page will not be graded.
You may assume throughout this exam that the code shown in Figure 1 of the booklet of Figures 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.

1. Female glow-worms attract males by glowing with part of their abdomen (called the "lantern"). Researchers believe the brightness of glow might be related positively to mating success. They measure the brightness of glow by the length of the lantern (in mm), and the mating success by the number of eggs laid by the female. Some of the data are shown in Figure 2, with some summary statistics below that.

   (a) [2] Based on Figure 3, do you think the researchers' belief is supported by the data? Explain briefly.

   (b) [2] Some predictions are shown in Figure 4, with the code that produced them at the top of the Figure. What precisely do the confidence limits in the first row of predictions tell you?

   (c) [3] Which one of the intervals in Figure 4 is shortest? Why does that make sense?

   (d) [3] Why does it make sense that the first interval in Figure 4 is shorter than the first interval in Figure 5?

2. Over time, a hospital admitted 360 patients with suspected heart attacks. Some of the patients had actually had a heart attack, but some had not (they had had something that looked like a heart attack, but was actually something else). A doctor believes that whether a suspected heart attack actually is one might depend on the level of creatinine kinase (something that can be measured with a blood test). However, creatinine kinase level can only be measured to a certain degree of accuracy, and so in a sample like this one, you would expect several patients to have the same creatinine kinase level. The data are shown in Figure 6. The three columns are:

   - `mck`: the measured creatinine kinase level
   - `ha`: the number of patients with that creatinine kinase level whose suspected heart attack actually *was* a heart attack
   - `nha`: the number of patients with that creatinine kinase level whose suspected heart attack actually *was not* a heart attack.

   (a) [2] Some code and the output from the code is shown in Figure 7. Why is it necessary to do this in preparation for the analysis that follows?

   (b) [2] A model and its output are shown in Figure 8. Does creatinine kinase level predict whether or not a suspected heart attack actually is one, and if so, is it a higher or lower value that tends to go with an actual heart attack? Explain briefly.

   (c) [2] A plot of predictions from the model is shown in Figure 9, with the code used to produce the plot above it. What are two ways in which this plot supports your conclusions from the previous part (or contradicts them, if that's what you think it does)?

   (d) [1] Some more predictions are shown in Figure 10, with the code that produced these predictions shown at the top of the Figure. What are these predictions of?

(e) [2] How are the predictions in Figure 10 consistent with one (or two) of the numbers in Figure 8? Explain briefly.

3. What determines whether a man is judged attractive by female college students? 38 men were rated for attractiveness, on a scale from A (most attractive) to D (least attractive). Each of the men had two other measurements taken:

   - `MaxGripStrength`: the man gripped a handheld dynamometer in their dominant hand and squeezed as hard as they could. The maximum of three grip strength measurements was taken (measured in kilograms).
   - `SHR`: shoulder to hip ratio; the circumference of the shoulders divided by the circumference of the waist.

   Some of the data are shown in Figure 11.

   (a) [2] A model is fitted, as shown in Figure 12. Why did I use `polr` (from package `MASS`) rather than something else?

   (b) [2] In the model shown in Figure 12, why do you think I added a squared term in `SHR`?

   (c) [2] Some predictions are shown in Figure 14. In the code above the predictions, why did I decide to use `pivot_wider`?

(d) [3] Look at Figures 13 and 14. How are they telling a consistent story about how `MaxGripStrength` influences attractiveness? Explain briefly.

(e) [3] According to Figure 14, how would you describe the effect of `SHR` on attractiveness? Explain briefly.

4. 90 males diagnosed with cancer of the larynx (where the vocal cords are) at a Dutch hospital took part in a study. Cancer cases are classified into one of four stages, numbered 1 through 4, in the column `stage` of this dataset. Stage 1 is the least advanced stage of the cancer, and Stage 4 is the most advanced, which would be expected to be worst. The numbers serve only to identify the stages; the numbers 1 through 4 have no meaning as numbers. The main aim of this study was to investigate the effect of the stage of cancer on survival times. The researchers also recorded the `age` of each patient, and the year of diagnosis (in `diagyr`, as two digits, the year minus 1900). The `time` in months from diagnosis until death, or until the end of the study, was also recorded, along with an indication `delta` of whether the patient was alive or dead at that time.

   Some of the data are shown in Figure 15.

   (a) [3] What code would I use to create a suitable response variable `y` for a Cox proportional-hazards model? This could be either a new column in the dataframe, or a separate variable outside the dataframe.

(b) [2] The values of `y` for the first twenty observations are shown in Figure 16. Why do some of them have a `+` next to them? How do you know?

(c) [2] A proportional-hazards model is shown in Figure 17. Why did I include the `drop1` output in addition to the `summary` output?

(d) [2] A second proportional hazards model is shown in Figure 18, and some further analysis is shown in Figure 19. Why did I need to do the further analysis, and what do you conclude from it?

(e) [4] Figure 20 shows some predicted survival curves. What do you conclude from the graph, and how is this conclusion consistent with the appropriate one of Figures 17 or 18? (Say which one of these last two Figures you are looking at.)

5. After you come out of a swimming pool, your fingers are wrinkled because they are wet. If you have to do a precision task with a wet object, is it better if your fingers are wrinkled or dry? To find out, 80 participants were observed doing a "transfer task" under various conditions. The task was to pick up an item with the right thumb and index finger, pass the item through a small hole, grab it with the left thumb and index finger, then put the item into a box through a hole in the lid. Each participant was timed, and the time to complete the whole task was recorded. Some of the data are shown in Figure 21. The columns of interest are:

   - `Time`: total time to complete the task, in seconds.
   - `Fingers`: whether the participant's fingers were `wrinkled` or `non` (not wrinkled)
   - `Objects`: whether the object being handled was `wet` or `dry`.

   For this question, carry out all tests at $\alpha = 0.10$.

   (a) [2] A plot is shown in Figure 22. What is the main thing that you learn from this plot? Explain briefly.

   (b) [2] Some analysis is shown in Figure 23. Is the result of this analysis what you were expecting, given your conclusion from the previous part? Explain (very) briefly, keeping in mind the α we are using.

   (c) [2] Someone tells you that the P-value for `Objects` in Figure 23 is the smallest, so you should now do a Tukey analysis of the types of objects. Do you think they are correct, or not? Explain briefly.

   (d) [3] Some more analysis is shown in Figure 24. What precisely does this enable you to conclude about the effects of `Fingers` or `Objects` or their combination? Looking back at Figure 22, summarize your conclusions clearly.

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