Problem Set 8

Statistical Methods In Engineering And Science

Due Date: 10:00 PM, May 26, 2023
Last Update: May 19, 2023

Prof. Alexander Giessing
Spring Quarter, 2023

Study Group: _____ _____ _____

Please upload your solution in a single pdf file on Canvas. Include all calculations, R-code, and figures (if applicable). All data sets are available on Canvas https://canvas.uw.edu/courses/1584511.

Question 1. We continue the analysis of the `olympic1500m.txt` data set from Problem Set 7. But now we analyze the problem as one of constructing a confidence interval for a proportion.

(a) Let X denote the number of races (out of 23 total races) won by a skater starting in the outer lane. If there is no advantage of starting on the outer lane, then it is reasonable to model X as having a $Bin(23, p)$ distribution with $p = 0.5$. Explain why.

(b) Use the normal approximation to the binomial distribution to derive lower 95% and 98% CIs for the unknown proportion p. *(Hint: Lectures of Week 8, Part 2, Slides 16ff. You can use Matlab, Wolfram Alpha, ... to solve the inequality.)*

(c) Based on your 95% CI, are you 95% confident that skaters starting in the outer line are indeed faster? What is your answer based on the 98% CI?

Question 2. Let $X_1, X_2, \ldots, X_n$ be a random sample from the uniform distr. on the interval $[0, \theta]$.

(a) Show that $U = \max_{1 \le i \le n} X_i/\theta$ has pdf

$$f_U(u) = \begin{cases} nu^{n-1} & 0 \le u \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(Hint: Lectures of Week 5, Part 1.)

(b) Use $f_U(u)$ to verify that

$$P\left(\alpha^{1/n} < \max_{1 \le i \le n} X_i/\theta \le 1\right) = 1 - \alpha,$$

and use this to derive a $100(1-\alpha)\%$ CI for θ.

(Hint: First, recall how to compute probabilities via densities. Second, for the CI adapt the steps of how to construct a CI for the mean, i.e. find suitable random variables L_n and U_n. See Lectures of Week 8, Part 1.)

Question 3. Briefly answer the following questions.

(a) Suppose that a popular sleeping pill makes 70% people sleep 7 hours or more. A new drug is introduced and claims to be more effective than the popularly used one.

(i) Formulate the null and alternative hypothesis.

(ii) What are the type I and type II errors in this context?

(b) A manufacturer of 40-amp fuses wants to make sure the mean amperage at which its fuses burn out is indeed 40.

(i) What null and alternative hypothesis would be of interest to the manufacturer?

(ii) What are the type I and type II errors in this context?

(c) True of False? (No justification required.)

(i) P-values are the chance that the null hypothesis is true.

(ii) When the p-value is 0.005, the null hypothesis looks plausible.

(iii) When the p-value is 10%, the null hypothesis cannot be rejected.

Question 4. Using carbon-14 dating we obtain repeated measurements of the age of a fossil. The sample average of 9 measurements is 2,050 years with a sample standard deviation of 90 years.

(a) Is there substantial evidence that the fossil is at least 2000 years old? Formulate the null and alternative hypothesis.

(b) Suppose that the measurements are normally distributed. Test the null hypothesis at significance level 5% using the appropriate test.

(c) What is the lower 95% CI for the expected age of the fossil? What can you conclude for the question in part (b) based on the result here?

Question 5. We return to the `olympic1500m.txt` data set one more time.

(a) Do speed skaters have an advantage in the 1500m race if they start in the outer lane? Formulate null and alternative hypothesis.

(b) In Question 4(b) on Problem Set 7 you (should) have established that the data are normally distributed. Use this information to test the null hypothesis at significance levels $\alpha = 5\%$ and 2% using the appropriate test.