Homework 9: MLEs

For each problem, remember you must briefly explain/justify how you obtained your answer, as correct answers without an explanation will not receive full credit. Moreover, in the event of an incorrect answer, we can still try to give you partial credit based on the explanation you provide.

In general, your goal in an explanation is to write enough that a student from class who has attended lecture, but not read the problem yet, could understand your approach, verify your reasoning, and believe your answer is correct. While we do not usually need to see arithmetic, you must include enough work that in principle one could rederive your answer with only a scientific calculator. For each problem, make sure to explicitly define all random variables you use, and be clear about how they are related to each other using proper notation (conditionals, summations, etc.).

Unless a problem states otherwise, you should leave your answer in terms of factorials, combinations, etc., for instance 26^7 or $26!/7!$ or $26 \cdot \binom{26}{7}$ are all good forms for final answers.

Instructions as to how to upload your solutions to gradescope are on the course web page.

Remember that you must tag your written problems on Gradescope.

Submission: You must upload a **pdf** of your written solutions to Gradescope under "HW 9". (Instructions as to how to upload your solutions to gradescope are on the course web page.) The use of latex is *highly recommended*. (Note that if you want to hand-write your solutions, you'll need to scan them. We will take off points for hand-written solutions that are difficult to read due to poor handwriting and neatness.)

Due Date: This assignment is due at 11:59 PM Wednesday June 2 (Seattle time, i.e. GMT-7).
You will submit the written problems as a PDF to gradescope. Please put each numbered problem on its own page of the pdf (this will make selecting pages easier when you submit), and ensure that your pdfs are oriented correctly (e.g. not upside-down or sideways). The coding problem will also be submitted to gradescope.

Collaboration: Please read the full collaboration policy. If you work with others (and you should!), you must still write up your solution independently and name all of your collaborators somewhere on your assignment.

For calculations that require evaluating integrals (unless we indicate otherwise), you must

(a) Show the integral to evaluate (e.g., $\int_0^2 z \cdot 2dz$)

(b) Show an antiderivative and the values to evaluate at (e.g., $z^2|_0^2$)

(c) Plug in the values and simplify (e.g., $2^2 - 0^2 = 4$)

1. MLE-1 [18 points]

Suppose you have a density

$$f_X(x) = \begin{cases} \frac{-|x|}{\theta^2} + \frac{1}{\theta} & \text{if } -\theta \leq x \leq \theta \\ 0 & \text{otherwise} \end{cases}$$

Let $x_1, x_2, \ldots, x_n$ be independent draws of a random variable from this distribution for some θ.

For this problem, *starting with part (b)* all of your work must be understandable without reference to any calculator (including wolframalpha). You may check your answers using calculators, but your explanation may not rely on them.

(a) Graph this density for a few values of θ. describe in a sentence or two what the shape is. The main goal of this part is for you to get intuition about the problem, you don't need to include the graphs [2 points]

(b) Write down the likelihood function for this problem; be careful. You should probably come back to this question after doing parts c and d to make sure you've handled edge cases. [3 points]

(c) Suppose you look at a value of θ such that $|x_1| \geq \theta$. What does the likelihood become in this case? Why? [3 points]

(d) In order to finish the problem without brutal algebra, we're going to assume that we've gotten exactly two independent samples x_1, x_2 and, moreover, assume that $x_1 = -x_2$.[1] Write the likelihood under this extra assumption. [3 points]

(e) Find the maximum likelihood estimator $\hat{\theta}$ under our extra assumptions. In taking a derivative, be sure to think about your answer in (c). You may skip the step of using the second derivative test to verify that your critical point is a maximizer (but you still must consider the piecewise part of the definition). [7 points]

2. MLE-2 [15 points]

In the final round of a rap battle, every member of the audience partakes in helping decide the winner between 3 rap battle contestants: Anna, Robbie and Kushal. Suppose that independently each member of the audience votes for Anna with probability θ_A, for Robbie with probability θ_R and for Kushal with probability $1 - \theta_A - \theta_R$. (Thus, $0 \leq \theta_A + \theta_R \leq 1$.) The parameters θ_A, θ_R are unknown. Suppose that $x_1, \ldots, x_n$ are n independent, identically distributed samples from this distribution. (Let n_A = number of x_i s equal to Anna (votes for Anna), let n_R = number of x_i s equal to Robbie (votes for Robbie), and let n_K = number of x_i s equal to Kushal (votes for Kushal).) What are the maximum likelihood estimates for θ_A and θ_R in terms of n_A, n_R, and n_K? In doing this problem, you do not need to do a second-derivative test (or any other test) to confirm you have a maximizer. You may assume any critical point you find is a maximizer

For this problem, all of your work must be understandable without reference to any calculator (including wolframalpha). You may check your answers using calculators, but your explanation may not rely on them.

3. Practice Exam [12 points]

We have put an old practice exam on the resources tab on Ed (or click here). This question asks you to take it under exam conditions and have you reflect on how it went. The goal of this exercise is to help you study for the final, and to frame any future studying you may want to do.

The practice exam was originally written as preparation for an in-person, closed-book exam. Your exam will look different as a result (e.g., we likely won't provide a reference sheet, since you can access everything you need; and we likely won't have as many T/F questions) As a result, we recommend you attempt the exam as though it were an individual, closed-book, in-person one. Specifically:

- Don't open the practice exam until you're ready to take it. Answer the exams on paper, working alone. (You will not submit the)
- Set a timer for 2 hours, and work for that time (or until you finish). When the timer goes off STOP, even if you aren't finished.
- Take the exam alone and on paper, without referencing anything other than the reference sheet (even though your exam is open-resource).
- After the 2 hours ends, compare your answers to the solutions (also on Ed).

Once you're done, answer the following questions. 2-3 sentences each should suffice for b-f (and the few words for part a).

(a) How did you do? For questions 1,2,3,4,5 say it was "completely right, mostly right, mostly wrong, didn't start"

(b) What topics did you struggle with? Are there any you think you should focus on for studying?

[1] This is not a normal assumption to make for an MLE calculation, but the algebra is much worse in the general case, and this very simplified case will still be good practice.

(c) What topics did you do better with? Are there any you think you don't need to study at all?

(d) Are there any topics that aren't covered that you were surprised not to see?

(e) How do you think a collaborative, take-home exam will be different? Will that affect what topics you're concerned about?

(f) Any other takeaways you haven't mentioned yet?