APPLIED MATH 115 HOMEWORK 2

This homework covers material in week three and four of the class. You are allowed to work in groups of up to three people from the class on this homework – every individual should submit their homework but please list your group members, and write a summary of 'author contributions' to the homework that everyone in the group agrees upon.

The learning goals of homework is several-fold:

(1) Start planning the first midterm project.
(2) Understand the extreme value distribution, and study empirically how robust this distribution to the underlying probability distribution – including validating empirically convergence to a universal distribution.
(3) Apply the extreme value distribution to two actual examples of rare events, one that we provide and another that you find yourself.
(4) Understand Keller's model of energy expenditure when running a race. Use this to construct optimal strategies for different running races.
(5) Understand the random walk and its relationship to the diffusion equation. Quantitatively compare simulations of random walkers to solutions of diffusion equation.
(6) Analyze the data you took in the molten chocolate cake lab.

Date: February 8, 2024.

Problem 1: *Midterm Project*

The first midterm project is due the week of March 5th. You should work in groups of 3 people, and study a question that is inspired from the first part of this class. (It might build on a topic that we have discussed, or it might use the methods and code we have discussed to address a different type of modeling problem.) The deliverables are

(1) A 3 minute presentation on what you have done, to be given to the class with your group on March 5 and 7.
(2) A brief report (4-6 pages long) that summarizes what and how you did it. Each group must turn in a single report. We will give both an example of the format we are looking for as well as a grading rubric.

For this problem set, we would like you to give an initial plan for your project. Please submit a PDF document to canvas answering the following questions **for each group** by February 16 at 5pm. Michael and Francesco will then go through and provide feedback via canvas on the projects in an effort to help you focus.

(1) What is the topic of your project? What question would you like to answer?
(2) Who are your group members?
(3) What methods or ideas will you use from the first part of the class?
(4) Do you have questions about how to start that we can help with?

Problem 2: *Extreme Value Distributions*

In class, we discussed the exteme value distribution and showed that for the Gaussian distribution and the exponential distribution, showing that the distributions had the same shape (modulo mean and variance).

(1) Redo this calculation computationally, following the class/section derivation for the powerlaw distribution with a heavy tail (use scipy.stats.powerlaw with $a < 1$ as mentioned in the Lecture notebook) distribution as well as any other distribution you choose to use that is a method in np.random. Determine whether the shape of the distribution has the same universality property we described in class.
(2) Let's now apply the extreme value distribution to actual data. We are providing data about the distribution of running times in the 100 meter dash. Compare the distribution of fastest times to the extreme value distribution. Does it agree? If not, could you hypothesize what the issue might be?
(3) Now look at the racing times for Usain Bolt. Can you make any statements about whether his times are consistent with the extreme value distribution?
(4) Find another example of a dataset where extreme values occur, and compare it to the extreme value distribution. A prize goes to the most creative solution!

Problem 3: *McMahon's Rowers, revisited*

In class we discussed McMahon's argument for rowers, which is an example of intelligent estimation. We are providing a dataset of rowing speeds in a python notebook taken from the 2022 and 2023 world championship.

(1) Please analyze this data and decide whether it obeys the $N^{1/9}$ law from McMahon.
(2) List factors in a rowing race that could mess up the basic argument.

Problem 4: *Estimation*

Let's try a estimation questions for yourself. Your goal here is to invent an intelligent estimate to come up with an answer based on facts that you might know.

(1) When if ever will Facebook contain more dead people than alive people?
(2) How many airplanes are currently in the sky?

Problem 5 *Random Walks and Diffusion*

(1) Let's consider a random walker with probability $p = 1/4$ of moving to the right and $q = 1/3$ of moving to the left. Since $p + q \neq 1$ there is some probability that the walker stands still. For definiteness, lets suppose that the walkers are at uniform density for $|x| < 1$, and that the walkers live on the real line $-\infty \leq x \leq \infty$.
   (a) Closely following the class discussion, derive the partial differential equation that describes the cloud of walkers. What is the diffusion constant and what is the advection speed?
   (b) Simulate the cloud of walkers. Compare the results of the simulation *quantitatively* to the solution to the diffusion equation we described in class. In particular, show that at long times, the derived PDE accurately and quantitatively reflects the solution.

(2) This week we carried out the molten chocolate cake lab (!). Work out the questions at the end of the lab, and in particular examine the data for L and the temperature, and estimate the transition temperature of cake batter is. Make sure to look at several values of L and take an average to get a good estimate.

(3) Use the calculator to come up with a reasonable guess at solutions to the diffusion equation that match your measured crust thickness L. Make a hypothesis for why the crust thickness is not uniform around the cake, and present evidence from this calculator to support or invalidate your hypothesis.

Advanced (Extra Credit) Problem *Keller's Model of Racing*

Note that this problem is not required – it is an optional extra credit problem for those who are interested in applying more advanced methodology. The goal here is to apply Keller's model of racing we discussed in class to actual data. Let's use the parameters maximum acceleration $P_{max} = 12m/sec^2$, resistance coefficient $k = 1.1$ per sec, initial energy reserve $E_0 = 2400m^2/s^2$, Aerobic energy distribution $\sigma = 40m^2/s^3$.

(1) Using the code from class, solve the optimal control problem with the parameters given to figure out hte time it would take for a world champion to run 200 meters, 400 meters, 1500 meters, 2000 meters, 3000 meters and 10000 meters.

(2) Divide the total distance into small enough segments that you can do the calculation in a reasonable amount of time in a colab. Start with 4 meter intervals and then try 2 meters. Comment on how the result changes when you change the time of the segment you divide the distance into. What is the appropriate segment length to use?

(3) Compare the results with the actual world records. You can find the world records at this link.

(4) Do you think this model is applicable when the racing distances get longer? What other factors might be important?

(5) Besides the time of the race, you will also find the optimal strategy for runners. As best as you can compare this to what actually happens with runners. Do the results make sense? Is there any part of the result that you disagree with? Why, and how would you fix the model?