CS114. Assignment 3. Final Problems.

Instructions

- Complete all the problems below and show your work.
- In addition to the LOs listed next to each problem, you will be graded on #MathTools, #CompTools, and #Professionalism for the overall quality of your text, math, and code.
- You may upload one Python notebook containing all your work, or a PDF with your text **and** math and a Python notebook with your code separately. No handwritten submissions.

Problem 1. Extreme Temperatures (#Probability, #ModelSelection)

Let $X_1, X_2, \ldots, X_{100}$ be the annual average temperature in Berlin in the years 2001, 2002, ..., 2100, respectively. Assume that average annual temperatures are sampled i.i.d. from a continuous distribution.

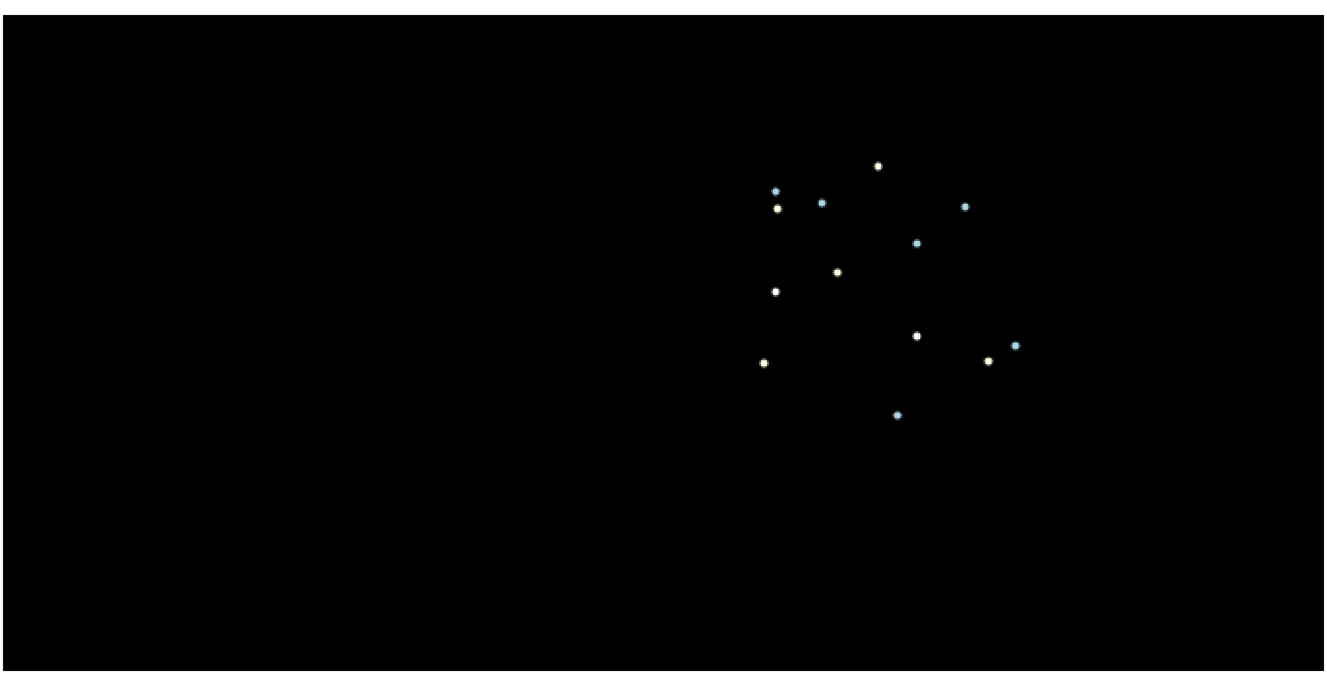
(Note: For this problem, we assume the temperature distribution doesn't change over time. With global warming, this is not a good assumption.)

A year is a record high if its average temperature is greater than those in all previous years (starting with 2001), and a record low if its average temperature is lower than those in all previous years. By definition, the year 2001 is both a record high and a record low.

1. In the 21st century (the years 2001 through 2100, inclusive), find the expected number of years that are either a record high or a record low.
2. Let N be an r.v. representing the number of years required to get a new record high after the year 2001. Find $P(N > n)$ for all positive integers n, and use this to find the PMF of N.
3. Write a short simulation to check your answers to parts (1) and (2).
4. Explain how you could use this model to determine whether or not global warming is really happening.

Problem 2. Starry Night (#ParameterEstimation)

It is a dark, clear night. A man sits inside his house and gazes at the stars through a square window. This is what he sees. What is the size of the window?



Assumptions:

- You cannot see the edge of the window since it is completely dark except for the stars outside.
- The window is perfectly square.
- The stars are points in the sky (they have negligible radius).
- The stars are distributed randomly, uniformly, and independently in the night sky.

Guidelines:

- The size (width and height) of the window is s and the center of the window is at (x, y). You need to infer the value of s but you should also think about how x and y play a role.
- You need to provide a distribution for s. A point estimate is not enough.

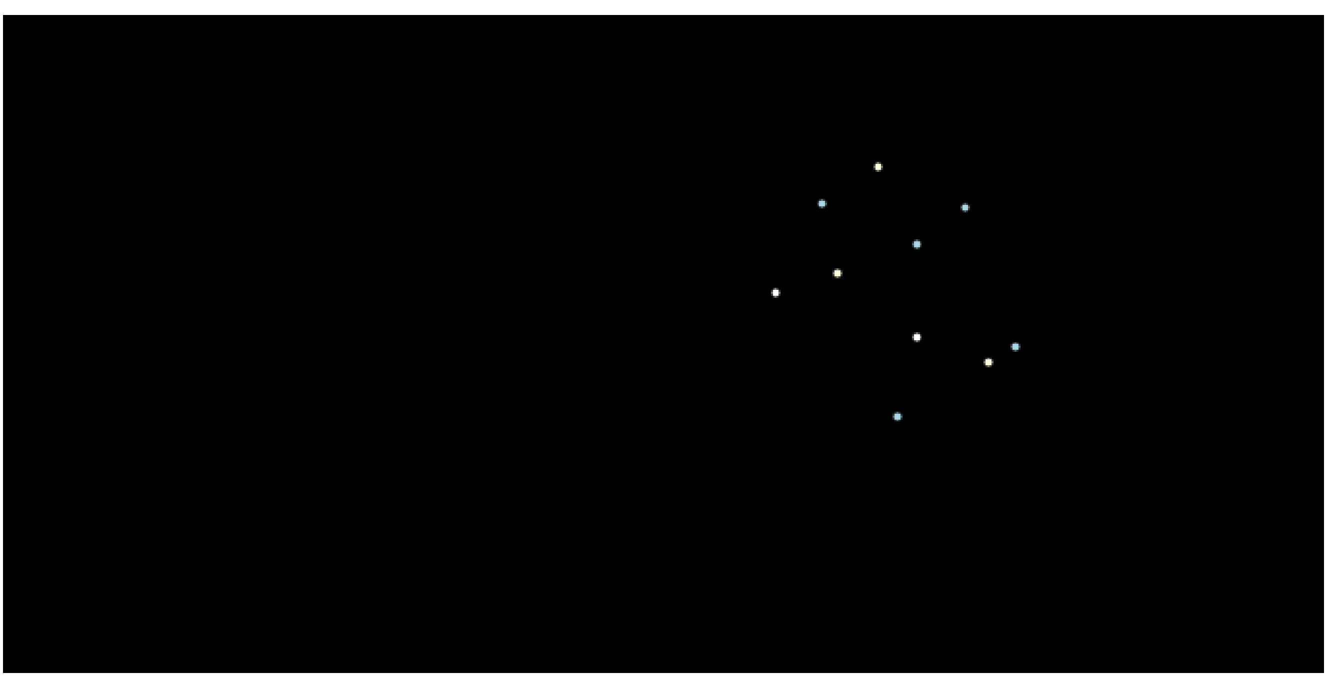
Data: The coordinates of the stars are given below. Your estimate for s must be in the same coordinate system as the data.

```
[[5.079, 1.493], [4.341, 1.657], [4.777, 1.811], [4.717, 1.280], [4.484, 1.936],
 [4.995, 1.444], [4.924, 1.922], [4.530, 1.720], [4.774, 1.525], [4.658, 2.049],
 [4.302, 1.439], [4.341, 1.971], [4.348, 1.919]]
```

Stretch Goal

(Changes to the story are marked in **bold**.)

It is a dark, clear night. A **hobbit** sits inside his house and gazes at the stars through a **circular** window. This is what he sees. What is the **radius** of the window?



Tasks:

- Use the coordinates of the first 10 stars above for this problem and disregard the last 3 stars completely — they are not visible to the hobbit.
- Carefully explain the differences in solving this problem for square and circular windows and how these differences impact the mathematics required for the problem.
- Solve the problem, providing a posterior distribution over the radius and the center of the circular window.
- If you make any approximations, be sure to explain what they are and how they might impact the accuracy of your final result.

The stretch goal is entirely optional — there is no penalty for not attempting it and there is no penalty for getting it wrong if you choose to try it. If you provide a **great, well-written solution** to this problem, you will get an **extra score of ⑤** on #ParameterEstimation. If you fall short of a ⑤, you can still earn a ④ for a slightly flawed solution but the errors have to be minimal. You cannot score less than ④ on this problem — instead you will get no extra grade if the attempt is insufficient. This is a hard problem and it is possible that nobody will get it right.