EE5311 Assignment 2 (AY24/25 Semester 2)

This assignment counts toward 25% assessment for the course. This assignment is due for submission by **20 April 2025**.

The assignment problem below is open-ended and may admit many possible solutions. Bearing in mind that the assignment provides you an opportunity to practice what you learn in this course, it is recommended that you think about how probabilistic programming techniques might help you solve the problem.

You **may** discuss the assignment problem and solution methodologies with other students and the instructor, but you are **not allowed** to share code or project report with other students. The Canvas discussion forum is a great place for discussions and/or clarification regarding this assignment.

> 🔥 Caution
>
> **Any copying or plagiarism will hamper your own learning, and may also result in disciplinary action against you**.

Problem Statement

A squash club keeps careful records of games played by the top ranked players at the club. The file `data2.csv` contains the results of games played by the top 10 women players at the club over the past 3 years. Each row in the file contains the names of the two players who played the game and the name of the winner. Assuming that the data is representative of the current skill levels of the players, we would like to predict the outcome of a series of 50 games to be played between Gloria and Ingrid over the next month. How many of the 50 games do you predict Gloria will win?

A simple model used to predict the outcome of a game between two players assumes that each player i has an intrinsic skill level μ_i and performance variability σ_i. For 95% of women squash players, the intrinsic skill level is between 10 and 30, and the performance variability is between 1 and 9. During a given game, the actual performance of player i is a random variable drawn from a normal distribution with mean μ_i and standard deviation σ_i. The player with the higher performance during that game wins. The performance in each game is independent of other games played. The skill and variability parameters for each player have

to be determined from historical data, assuming that they are constant over the time period covered by the data.

Submission requirements

(25 marks)

You are required to submit a `zip` file (name the file with your login ID, e.g., `e1234567.zip`) containing the following:

- Brief report (**page limit: 3 pages**) in `pdf` format (please do NOT submit Word documents!), clearly outlining the method used and an estimate of the number of games out of 50 that Gloria is predicted to win. State any assumptions you make and justify them. Comment on the reliability of your prediction.
- Well-commented runnable code yielding your results. Running this code should create the results shown in your report. You should include a `README.txt` file with instructions on how to run your code.