Midterm **Name:** ____________________
Mar 10, 2023, 12:30-2:20pm **Student ID:** ____________________
CMPT 727: Spring 2023
Instructor: Maxwell Libbrecht

Instructions:
This exam is open-book and open-internet. You may refer to any written or electronic materials, and you may use an internet-connected device to view the web (e.g. the textbook web page, Google, Wikipedia, etc). No two-way communication is allowed (email, text messages, Discord, etc).

You can use any legible format that supports mathematics for your solution, including paper, a stylus on a tablet, or LaTeX. If you use a text editor such as Word, you must use the formula editor (not plain text).

After you finish, please attest the following in your solution: "I have had no communication with anyone other than the instructors during this exam."

Problem 1. Suppose N sports teams play each other in a tournament. We imagine that each team's ability can be expressed as a number a_i. When teams i and j play against each other, we define the outcome as $w_{ij} = \begin{cases} 1 & \epsilon_{ij} \geq 0 \\ 0 & \epsilon_{ij} < 0 \end{cases}$, where $\epsilon_{ij} \sim \mathcal{N}(a_i - a_j, \sigma^2)$. Every team has played each other exactly once. We would like to estimate the teams' abilities $a_{1:N}$ based on their win/loss record.

Note: For all questions, you can use the Gaussian CDF function $\Phi(x; \mu, \sigma) = P(y < x)$ if $y \sim \mathcal{N}(\mu, \sigma)$ in your solution (\Phi in LateX). Your answers should not include any unsolved integrals, derivatives or probability expressions (i.e. statements of the form "$P(...)$").

(a) Draw a model of this problem using plate notation. Include a, w, and σ as random variables in your model.

(b) Suppose (for this question only) that we observe all ϵ_{ij}. Prove that

$$P(a_1 | a_{2..N}, \epsilon) = \mathcal{N}(\mu_1, \sigma_1^2),$$
$$\text{where} \quad \mu_1 = \frac{1}{N-1} \sum_{i=2}^{N} (\epsilon_{1i} + a_i),$$
$$\sigma_1^2 = \frac{\sigma^2}{N-1}.$$

Assume a uniform prior on a_1, i.e. $a_1 \sim \mathcal{N}(0, \inf)$. Hint: Apply Bayes' rule for Gaussians with $Z = a_1$.

(c) Write down the log-likelihood for learning $a_{1..N}$ given $w_{1:N,1:N}$.

(d) Suppose that we constraint team's ability $a_i \in [0, 1]$, $\forall i \in \{1..N\}$. Write down the Lagrangian formula to optimize the log-likelihood subject to the constraint on a. (You do not need to solve it).

(e) Suppose you would like to use stochastic gradient descent to learn the MLE of $a_{1..N}$ given the w_{ij}'s. Derive an update formula for a stochastic optimization step for $a_{1..N}$ given a single w_{ij}. Include a learning rate parameter η in your formula.

(f) Suppose we estimate $a_1^{(1)} = a_2^{(1)} = 0.5$ and we observe $w_{12} = 1$. Using your answer to the previous question, calculate our updated estimate $a_1^{(2)}$ given $\eta_1 = 0.5$. Assume all Lagrangian multipliers equal 0. You may either express your answer in a way that could be plugged into a calculator or find a numerical answer yourself, using e.g. `scipy.stats.norm.cdf(x)`.