EE5311 Differentiable and Probabilistic Computing: Assignment 2
XU NUO, A0313771H

1 Introduction

1.1 Objective

A squash club has recorded the game results among 10 players over the past three years, which are stored in the file "data2.csv". Each entry in the dataset contains the names of the two players involved in a game and the winner of that game. Given this historical data, the objective is to predict how many times Gloria is expected to win out of 50 future games against Ingrid, based on their past performance.

1.2 Assumptions

According to the statement of the problem, we make the following assumptions about the players and games:

- A player's actual performance in a game is assumed to follow a normal distribution determined by their intrinsic skill level μ and performance variability σ. The player with the higher performance wins the game.
- 95% of the players' μ values lie between 10 and 30, and their σ values lie between 1 and 9.
- Each player's performance distribution is considered stationary over time, meaning it does not vary across different time periods.
- All games are treated as independent and identically distributed events.
- The training data is assumed to be reliable and unbiased.

2 Methodology

2.1 Maximum Likelihood Estimation (MLE) with Monte Carlo sampling

Based on the problem statement, I model each player as a random variable, where their actual performance in each game follows a normal distribution:

$$x_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$$

with μ_i representing the intrinsic skill level of player i, σ_i representing the performance variability of player i. For a single game, if the winner's performance follows $x_w \sim \mathcal{N}(\mu_w, \sigma_w^2)$, the loser's performance follows $x_l \sim \mathcal{N}(\mu_l, \sigma_l^2)$, and the person with better performance wins, then it means the winner performed better than the loser:

$$P(winner\ wins) = P(x_w > x_l)$$

The difference between two normal distributions still satisfies a normal distribution, so the winning probability can be expressed as:

$$x_w - x_l \sim \mathcal{N}(\mu_w - \mu_l, \sigma_w^2 + \sigma_l^2)$$

$$P(x_w > x_l) = P(x_w - x_l > 0) = \Phi\left(\frac{\mu_w - \mu_l}{\sqrt{\sigma_w^2 + \sigma_l^2}}\right)$$

With Φ being cumulative distribution function of the standard normal distribution.
The given data has many games recorded, so we can construct a winner's odds for each game, which is transformed into a maximum log-likelihood problem:

$$logL = \sum_{All\ games} log\ \Phi\left(\frac{\mu_w - \mu_l}{\sqrt{\sigma_w^2 + \sigma_l^2}}\right)$$

We need to find a set of parameters μ_i and σ_i to maximize the positive likelihood, which means minimizing the negative log-likelihood. This task can be solved by function *scipy.optimize.minimize()*. After completing the parameter fitting, we obtain the estimated values of μ and σ for each player that best explain the historical game outcomes. These parameters define each player's potential performance range, which can be visualized using normal distribution curves.

Subsequently, the μ and σ values for Gloria and Ingrid are extracted to construct their respective normal distributions. A Monte Carlo simulation is then performed: in each trial, 50 performance samples are drawn from each player's distribution, and Gloria's number of wins is counted by comparing the sampled values. Repeating this process 1000 times yields a distribution of predicted win counts, which can be visualized as a histogram.

2.2 Bayesian Inference with Markov Chain Monte Carlo (MCMC) Sampling

As stated in the problem, 95% of the players' μ values lie between 10 and 30, and their σ values lie between 1 and 9. I further implement a Bayesian inference framework that estimates not only the most likely values of parameters μ and σ, but also their full posterior distributions. This provides a probabilistic understanding of the uncertainty associated with each player's skill and performance variability.

Prior distributions are placed on both μ_i and σ_i. Specifically, $\mu_i \sim TruncatedNormal(0,20,5)$, ensuring that the intrinsic skill level remains positive and centered around 20 and falls between 10 and 30 mostly. $\sigma_i \sim TruncatedNormal(0,5,2)$, ensuring that the performance variability remains positive and most of it falls between 1 and 9.

As in the MLE model, the probability that the winner of a game indeed outperformed the opponent is expressed as:

$$P(winner\ wins) = \Phi\left(\frac{\mu_w - \mu_l}{\sqrt{\sigma_w^2 + \sigma_l^2}}\right)$$

Using this expression, we define a Bernoulli likelihood over all games, where the observed outcome is always a win for the recorded winner.

The posterior distribution of the parameters is inferred using MCMC sampling, specifically the No-U-Turn Sampler (NUTS) provided by *NumPyro*. I run the sampler with 500 warm-up steps and collect 1000 samples from the posterior. These samples represent possible values of μ and σ for each player that are consistent with the observed game outcomes and the specified priors.

Similarly, a histogram is then plotted to visualize the predicted number of games Gloria is expected to win out of 50, repeated over 1000 posterior samples. Additionally, I compute and visualize the 95% credible intervals for each player's posterior μ and σ, highlighting both the estimated skill levels and the uncertainty associated with them.

3 Results and Discussion

The results of the estimation using these two methods are shown in the following figures. Fig. 2 and Fig. 3 present the predicted number of wins for Gloria in 50 future games against Ingrid, using the MLE and Bayesian approaches respectively. Both histograms show unimodal, bell-shaped distributions, suggesting that Gloria is more likely to win between **9 to 13** games, though exact numbers vary.

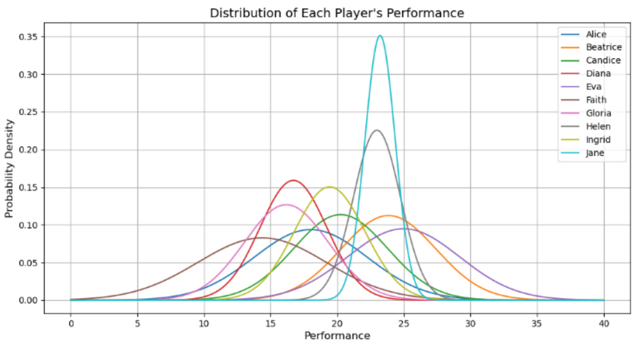
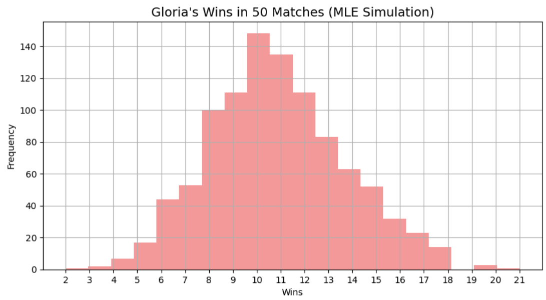


Fig. 1. Distribution of each player's performance (MLE)



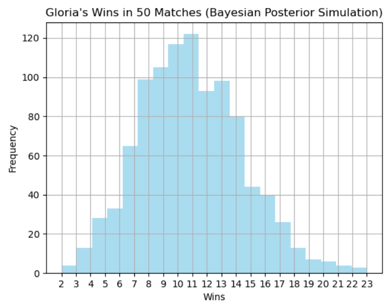


Fig. 2. Gloria wins in 50 matches (MLE)　　Fig. 3. Gloria wins in 50 matches (Bayesian)

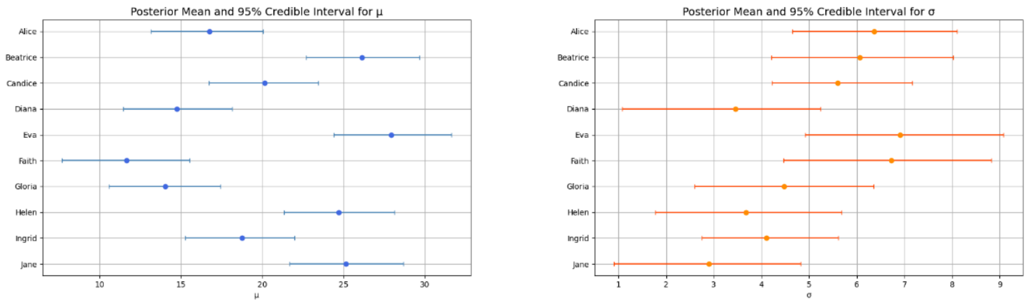


Fig. 4/5. Posterior mean and 95% credible interval for μ and σ (Bayesian)

As seen in Fig. 1, the MLE-based performance distributions suggest Ingrid is moderately more skilled than Gloria. Fig. 4/5 from the Bayesian model confirm this, with Gloria's posterior μ lower than that of Ingrid, and her σ higher than Ingrid's with credible interval spanning a higher range. These indicate that Gloria's intrinsic skill level is lower than Ingrid's, and her performance is relatively more unstable, which justifies the model's prediction that Gloria is less likely to win the majority of games.

A comparison between the MLE method and Bayesian inference shows that MLE aims to find a single set of parameters that maximizes the likelihood of the observed data, where both μ and σ are point estimates and the uncertainty in parameter estimation is ignored. In contrast, Bayesian inference combines prior knowledge with observed data to infer the posterior distributions of the parameters. The parameters are represented by a set of samples from the posterior, which provides a more realistic reflection of the underlying uncertainty.