

# HW6: Bayesian Inference

## Summary of the dataset:

Each row is a football match. The target variable is GoalsScored (= the number of goals scored by both teams combined for that match). Explanatory variables describe the two teams before the match start (H = home team, A = away team). Variables:

- ScoreRate = goals scored per game
- ConcedeRAte = goals conceded per game
- CornerRatio = what proportion of corners does the team have, on average, during their matches (0.5 = same as opponents)
- FoulRatio = what proportion of fouls does this team commit 100 matches are test obseervations (missing GoalsScored).

## Part 1 (Grades 6–7):

Provide insight into the relationship between the explanatory and GoalsScored variables with a Poisson GLM. Use Bayesian inference (you may use any library that supports MCMC inference and diagnostics). Study standard MCMC diagnostics (autocorrelation, traceplot, effective sample size, Rhat) and apply them to diagnose potential issues with MCMC computation.

## Part 2 (Grades 8–9):

Fit the model from PART 1 using Laplace Approximation. You may use libraries for general-purpose optimization and linear algebra, but the rest of the implementation of Laplace Approximation has to be your own. Validate your fit using MCMC samples from PART 1. For full credit, derive the gradient and Hessian by hand. Use the model to make point predictions for the test cases for three different loss functions:

- squared error
- absolute error
- accuracy (1 if we predict correct number of goals, 0 otherwise; for this loss function it makes sense to make only non-negative integer predictions instead of non-negative continuous).

## Part 3 (+0, +0.5, +1.0 Grade):

Polish your report (language, plots, tables, structure, everything counts).

## General

- In addition to these instructions, the homework includes a `football.csv` file.
- Submit a pdf report (all parts combined into a single pdf; no more than 1 page per part!)
- Feel free to use any tools, including LLMs and collaboration with others, but keep in mind that our goal is to understand what we are doing and not merely to do. Your work will be graded based on your understanding of your code, report, and the subject matter in general.