HW7: Bayesian Logistic Regression

Marko Ivanovski

Code: https://github.com/marcheivanovski/MLDS1 (folder Homework7)

Introduction

In this homework, we showcase how Bayesian inference works on logistic regression by applying it to a simple dataset.

Prior knowledge

What is your personal opinion about the coefficient beta for distance? Before building any models my intuition was that distance will negatively influence the target variable. The bigger the distance the less likely is the shot to be made. My prior opinion can be summarized by a normal distribution with some negative mean and a bigger variance just to be more general and uninformative. Regarding angle, I didn't think it influences the shot success by a lot. Therefore, my prior can be summarized by a normal distribution around zero (maybe a bit shifted towards the negative side).

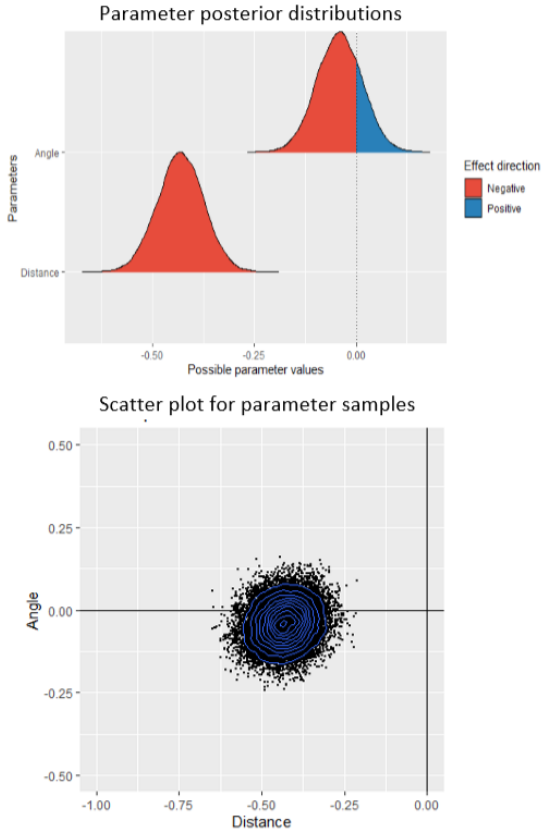

Figure 1. Posterior distribution, and scatter plot of LR coefficients on the whole dataset.

Working with the whole dataset

Before I started any analysis I standardized both features so that the comparison between posterior attribute distribution can be fair. Then, I generated a chain of 20000 samples. The posterior distribution for the model parameters and the scatter plot of the samples are in Figure 1. Default prior (N(0,2.5)) was used for all parameters. The ESS for intercept, angle and distance was: 18 176, 19 313, and 19 293 which meant our sampling was very effective.

Which is more important for shot success, angle or distance? To answer this question, I calculated the probability P($|\beta_1| - |\beta_2| > 0$)=0, ($\beta_1$ is angle and β_2 is distance). This means that when we compare pairwise magnitudes of both coefficient values, the distance coefficient will always be bigger hence more important.

Does shot success increase or decrease with increasing angle (the further on the sides we are)? To answer this question I integrated the posterior to 0 (angle impacts negatively) or from 0 (angle impacts positively). It turned out that P($|\beta_1| < 0$)=0.7621 or P($|\beta_1| > 0$)=0.2379. In conclusion, this parameter is more likely to be negative which means a bigger angle means a less likely successful shot.

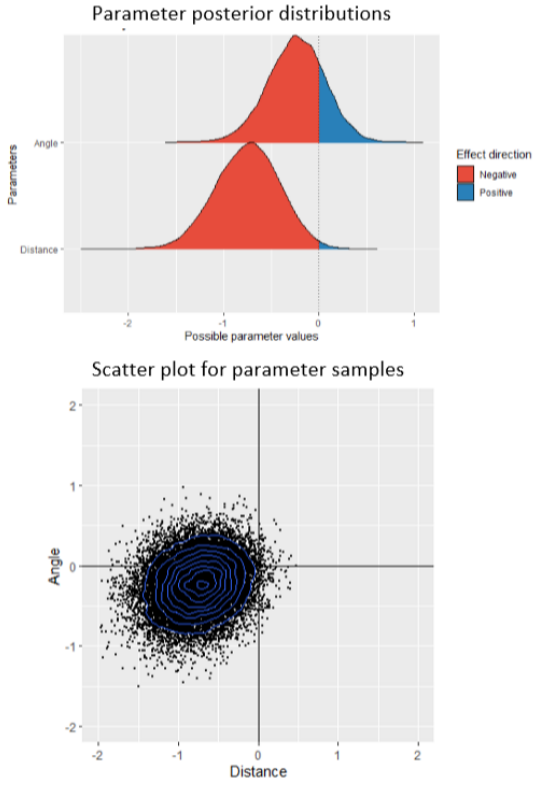

Figure 2. Posterior distribution, and scatter plot of LR coefficients on randomly chosen 50 instances from the dataset.

Working with 50 samples from the dataset

When working with a smaller amount of samples from the dataset the posterior is expect-idly wider. This gives information that the model is not as sure about the parameters as it was when it saw the whole dataset. Figure 2 shows the distance posterior is shifted to the right and the angle posterior is similar just wider. The ESS was also similarly high as previously.