Applied Bayesian Statistics
Practical 2
Metropolis-Hastings. Logistic regression. Modelling from first principles.

Robin Ryder

We consider the data `golf2.txt`. We wish to explain the success at golf putting, explained by the distance to the hole.

1. Explore the data and represent graphically the success frequence as a function of distance.

1 Logistic regression

2. We consider a logistic regression model:

$$P[Y_i = 1] = \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}}$$

   What are reasonable priors for β_0 and β_1?

3. Write the likelihood and posterior associated with this model. Is the posterior distribution easy to sample from?

4. Simumlate a pseudo-sample from this posterior via MCMC thanks to the `MCMClogit` function in package `MCMCpack`. Check convergence and adapt the algorithmic parameters as necessary.

5. Repeat with another prior and compare the results.

2 Modelling from first principles

6. The ball radius is $r = 0.07$ and the hole radius is $R = 0.177$ (all distance are in feet). Propose a model based on the geometric properties of the problem, assuming the player hits the ball with some random angle α.

7. Write the likelihood of your model, and choose a prior for any parameters.

8. Get a pseudo-sample of the posterior thanks to the function `MCMCmetrop1R`. Assess convergence.

9. Modify the algorithmic parameters (`tune`, `burnin`...) until the output is acceptable. What is the impact of the starting point `theta.init`?

10. Compare the data fit of both models.

11. Propose an extension of the second model, and check whether the fit is improved.