Bayesian Analysis on Rubik's Cube solving time

Ankit Tewari, Hong-Tan Lam, Cecilia Siliato Robles

May 2018

1 Abstract

Solving time of Rubik's cubes have been collected and analyzed. By having 4 different prior models on how the solving times are distributed, we try to select the best model using Bayes factor after the collection of the dataset. By using Markov chain Monte Carlo, we will try to predict how future solve-times are distributed, with the purpose of trying to help the Rubik's cube solver try to improve his/her times.

Contents

1 Abstract **1**

2 Introduction **2**

3 Models **2**
- 3.1 Model 1: $\{Y \sim N(\mu, 1.5), \quad \mu \sim \text{Uniform}(0, \infty)\}$ 2
- 3.2 Model 2: $\{Y \sim N(\mu, 1.5), \quad \mu \sim N(11, 0.8)\}$ 3
- 3.3 Model 3: $\{Y \sim N(\mu, 1.5), \quad \mu \sim Unif(9, 13)\}$ 5
- 3.4 Model 4: $\{Y \sim N(\mu, 1.5), \quad \mu \sim N(11, \hat{\sigma}), \quad \hat{\sigma} \sim Beta(8, 2)\}$ . . . 6

4 Experiment **7**

5 Model selection **7**

6 Results and Prediction **9**
- 6.1 Model 1 10
- 6.2 Model 3 11
- 6.3 Model 4 11
- 6.4 Model 2 12

7 Discussion **14**

8 Conclusion **15**

9 Code **15**

2 Introduction

Rubik's cube is a traditional mechanical puzzle that consists of a cube of six faces that in turn are composed of 9 six-color stickers. The goal consists of getting each of the 6 faces to contain its 9 stickers of the same color.
Currently it is considered the best selling toy in the world and in the last 10 years there have been hosted competitions all over the world where the goal of the competitors is to solve the cube in the shortest amount of time. Taking advantage of the fact that one of the members of the group is a Rubik's cube solver, we have decided to do a study about the time it takes to solve it.
To do this, we start from a database of 50 observations, where each observation contains the time (in seconds) it takes to solve the cube. The objective of this work is to make a model that adapts to the given data in order to predict the time it will take to solve the cube the next n times. For this, we will evaluate several possible Bayesian models and later we will decide which one will best suit our data.

3 Models

For our project, we have chosen 4 different type of models. We assume that the solving time is normally distributed (from prior experience), with an unknown mean and a standard deviation of 1.5. All 4 models will therefore have the same likelihood function. However, we are not so sure about how the mean is distributed. It might be uniformly distributed, normally distributed or perhaps it is dependent on another parameter (hyper parameter). Because of the different possibility of the mean's distribution, we propose 4 models:

1. $y \sim \mathrm{N}(\mu, \sigma)$
   $\mu \sim \mathrm{Uniform}(0, \infty)$
2. $y \sim \mathrm{N}(\mu, \sigma)$
   $\mu \sim \mathrm{N}(11, 0.8)$
3. $y \sim \mathrm{N}(\mu, \sigma)$
   $\mu \sim \mathrm{Uniform}(9, 13)$
4. $y \sim \mathrm{N}(\mu, \sigma)$
   $\mu \sim \mathrm{N}(11, \hat{\sigma})$
   $\hat{\sigma} \sim \mathrm{Beta}(8, 2)$

3.1 Model 1: $\{Y \sim N(\mu, 1.5), \quad \mu \sim \mathrm{Uniform}(0, \infty)\}$

We were curious about if having a prior with no information could give good results, as there is a possibility that our assumptions are really bad. The prior is therefore defined so that there is an equal probability of the mean being any possible positive value. Since the prior does not integrate to 1, it is both improper and non-informative.

- **Prior:**
  The prior is a non-informative prior, a uniform distribution with value 1 when μ is a non-negative number: $\Pi(\mu) = 1$ when $\mu \geq 0$. Otherwise $\Pi(\mu) = 0$.

- **Likelihood:**
  $$p(y|\mu) = \prod_{i=1}^{50} \left(\frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-\frac{1}{2}(\frac{y_i - \mu}{1.5})^2} \right) = \frac{1}{(2\pi)^{25} \cdot \sigma^{50}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{50} (y_i^2 - 2\mu y_i + \mu^2)},$$
  where $\sigma = 1.5$.

- **Posterior:**
  We calculate the posterior:
  $$p(\mu|y) = \frac{p(y|\mu)\Pi(\mu)}{\int_0^\infty p(y|\mu)\Pi(\mu)d\mu} = \frac{p(y|\mu)}{\int_0^\infty p(y|\mu)d\mu} = \frac{\frac{1}{(2\pi)^{25} \cdot \sigma^{50}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{50} (y_i^2 - 2\mu y_i + \mu^2)}}{\displaystyle\int_0^\infty \frac{1}{(2\pi)^{25} \cdot \sigma^{50}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{50} (y_i^2 - 2\mu y_i + \mu^2)} d\mu}$$

- **Prior predictive:**
  $$p(\tilde{y}) = \int_0^\infty p(\tilde{y}|\mu)\Pi(\mu)d\mu = \int_0^\infty \frac{1}{(2\pi)^{25} \cdot \sigma^{50}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{50} (\tilde{y}_i^2 - 2\mu \tilde{y}_i + \mu^2)} \cdot 1 d\mu$$

- **Posterior predictive:**
  $$p(\tilde{y}|y) = \int_0^\infty p(\tilde{y}|\mu)p(\mu|y)d\mu$$
  $$= \int_0^\infty \frac{1}{(2\pi)^{25} \cdot \sigma^{50}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{50} (\tilde{y}_i^2 - 2\mu \tilde{y}_i + \mu^2)} \cdot \frac{\frac{1}{(2\pi)^{25} \cdot \sigma^{50}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{50} (y_i^2 - 2\mu y_i + \mu^2)}}{\displaystyle\int_0^\infty \frac{1}{(2\pi)^{25} \cdot \sigma^{50}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{50} (y_i^2 - 2\mu y_i + \mu^2)} d\mu} d\mu$$

3.2 Model 2: $\{Y \sim N(\mu, 1.5), \quad \mu \sim N(11, 0.8)\}$

In this second model, we are trying to consider that the mean of the solving time is distributed as a normal with parameters $\mu_1 = 11$ and $\sigma_1 = 0.8$. We have considered that because before knowing the data we suppose that the mean is something around 11 and if we consider a large standard deviation, then our model will be less concrete.

Now we are going to see the different distributions of this model (prior, likelihood, posterior, prior predictive and posterior predictive):

- **Prior:**
As we have decided, in this model, we will take $\mu \sim N(11, 0.8)$, so the prior has the expression:

$$\pi(\mu) = \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}}$$

where $\mu_0 = 11$ and $\sigma_0 = 0.8$

- **Likelihood:**
The likelihood is the same as on the rest of models:

$$p(y|\mu) = \prod_{i=1}^{50} \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i-\mu)^2}{2\sigma^2}} \right) = \frac{1}{(2\pi)^{25} \cdot \sigma^{50}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{50} (y_i^2 - 2\mu y_i + \mu^2)}$$

where $\sigma = 1.5$

- **Posterior:**
In order to compute the posterior distribution, remember that we should follow the formula:

$$p(\mu|y) = \frac{p(y|\mu)\pi(\mu)}{\int_{-\infty}^{\infty} p(y|\mu)\pi(\mu) d\mu}$$

If we attend at the numerator:

$$p(y|\mu)\pi(\mu) = \frac{1}{(2\pi)^{25} \cdot \sigma^{50}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{50} (y_i^2 - 2\mu y_i + \mu^2)} \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}} =$$

$$= \frac{1}{(2\pi)^{\frac{51}{2}} \sqrt{\sigma_0^2 \sigma^{2n}}} exp \left\{ \frac{-\mu^2 + 2\mu\mu_0 - \mu_0^2}{2\sigma_0^2} - \sum_{i=1}^{50} \frac{y_i^2 - 2\mu y_i + \mu^2}{2\sigma^2} \right\}$$

So the expression is:

$$p(\mu|y) = \frac{\frac{1}{(2\pi)^{\frac{51}{2}} \sqrt{\sigma_0^2 \sigma^{2n}}} exp \left\{ \frac{-\mu^2 + 2\mu\mu_0 - \mu_0^2}{2\sigma_0^2} - \sum_{i=1}^{50} \frac{y_i^2 - 2\mu y_i + \mu_0^2}{2\sigma^2} \right\}}{\int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{51}{2}} \sqrt{\sigma_0^2 \sigma^{2n}}} exp \left\{ \frac{-\mu^2 + 2\mu\mu_0 - \mu_0^2}{2\sigma_0^2} - \sum_{i=1}^{50} \frac{y_i^2 - 2\mu y_i + \mu^2}{2\sigma^2} \right\} d\mu}$$

We should replace in this expression the given data and so, we will obtain the posterior value.

- **Prior predictive:**

$$p(\tilde{y}) = \int_{-\infty}^{\infty} p(\tilde{y}|\mu)\pi(\mu) d\mu = \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{25} \cdot \sigma^{50}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{50} (\tilde{y}_i^2 - 2\mu \tilde{y}_i + \mu^2)} \cdot \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{(y-\mu_0)^2}{2\sigma_0^2}} d\mu$$

- **Posterior predictive:**

$$p(\tilde{y}|y) = \int_{-\infty}^{\infty} p(\tilde{y}|\mu)p(\mu|y)d\mu = \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{25} \cdot \sigma^{50}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{50}(\tilde{y}_i^2 - 2\mu\tilde{y}_i + \mu^2)} \cdot$$

$$\cdot \frac{\frac{1}{(2\pi)^{\frac{51}{2}}\sqrt{\sigma_0^2\sigma^{2n}}} exp\left\{\frac{-\mu^2+2\mu\mu_0-\mu_0^2}{2\sigma_0^2} - \sum_{i=1}^{50}\frac{y_i^2-2\mu y_i+\mu^2}{2\sigma^2}\right\}}{\int_{-\infty}^{\infty}\frac{1}{(2\pi)^{\frac{51}{2}}\sqrt{\sigma_0^2\sigma^{2n}}} exp\left\{\frac{-\mu^2+2\mu\mu_0-\mu_0^2}{2\sigma_0^2} - \sum_{i=1}^{50}\frac{y_i^2-2\mu y_i+\mu^2}{2\sigma^2}\right\} d\mu} d\mu$$

3.3 Model 3: $\{Y \sim N(\mu, 1.5), \quad \mu \sim Unif(9, 13)\}$

In the 3rd model, we assume that the mean is uniformly distributed between 9 and 13 seconds. This prior is a bit similar to the first model (uniform) but it has a much smaller support. We therefore get a much more specific prior, where the specific interval was chosen because we believe that the mean is around 11 seconds.

- **Prior:**
  As we have decided, in this model, we will take a uniform prior with minimum value 9 and maximum value 13 i.e. $Unif(9, 13)$. So the prior has the expression:

$$\pi(\mu) = \begin{cases} \frac{1}{4}, & x \in [9, 13] \\ 0, & \text{otherwise} \end{cases}$$

- **Likelihood:**
  The likelihood is the same as on the rest of models:

$$p(y|\mu) = \prod_{i=1}^{50}\left(\frac{1}{\sqrt{2\pi}\cdot\sigma} e^{-\frac{1}{2}(\frac{y_i-\mu}{1.5})^2}\right) = \frac{1}{(2\pi)^{25}\cdot\sigma^{50}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{50}(y_i^2-2\mu y_i+\mu^2)}$$

  where $\sigma = 1.5$

- **Posterior:**
  In order to compute the posterior distribution, remember that we should follow the formula:

$$p(\mu|y) = \frac{p(y|\mu)\pi(\mu)}{\int_0^{\infty} p(y|\mu)\pi(\mu)d\mu}$$

  So the expression is:

$$p(\mu|y) = \frac{1}{(2\pi)^{25}\cdot\sigma^{50}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{50}(y_i^2-2\mu y_i+\mu^2)} \frac{1}{4}$$

$$p(\mu|y) = \frac{\frac{1}{(2\pi)^{25}\cdot\sigma^{50}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{50}(y_i^2-2\mu y_i+\mu^2)}\frac{1}{4}}{\displaystyle\int_9^{13} \frac{1}{(2\pi)^{25}\cdot\sigma^{50}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{50}(y_i^2-2\mu y_i+\mu^2)}\frac{1}{4}d\mu}$$

On simplification, the constant terms in numerator and denominator cancel out and we get,

$$p(\mu|y) = \frac{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{50}(y_i^2-2\mu y_i+\mu^2)}}{\displaystyle\int_9^{13} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{50}(y_i^2-2\mu y_i+\mu^2)}d\mu}$$

We should replace in this expression the given data and so, we will obtain the posterior value.

- **Prior predictive:**

$$p(\tilde{y}) = \int_{-\infty}^{\infty} p(\tilde{y}|\mu)\pi(\mu)d\mu = \int_9^{13} \frac{1}{(2\pi)^{25}\cdot\sigma^{50}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{50}(\tilde{y}_i^2-2\mu\tilde{y}_i+\mu^2)}\cdot\frac{1}{4}d\mu$$

- **Posterior predictive:**

$$p(\tilde{y}) = \int_{-\infty}^{\infty} p(\tilde{y}|\mu)\pi(\mu)d\mu$$

or, we can simplify and write it as,

$$p(\tilde{y}|y) = \int_9^{13} \frac{1}{(2\pi)^{25}\cdot\sigma^{50}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{50}(\tilde{y}_i^2-2\mu\tilde{y}_i+\mu^2)}. \frac{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{50}(y_i^2-2\mu y_i+\mu^2)}}{\displaystyle\int_9^{13} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{50}(y_i^2-2\mu y_i+\mu^2)}d\mu}d\mu$$

3.4 Model 4: $\{Y \sim N(\mu, 1.5), \quad \mu \sim N(11, \hat{\sigma}), \quad \hat{\sigma} \sim Beta(8, 2)\}$

There is a possibility that the mean depends on another parameter. We therefore decided to set the prior, such that the mean is normally distributed with mean 11 and a standard deviation $\hat{\sigma}$ that is beta-distributed. The parameters for the beta distribution was chosen such that the standard deviation would vary between 0.4 and 1. This is because we saw that with this range of standard deviation, we get a mean that varies nicely from 9 to 13, which is what we think is a nice interval for the mean.
We define

$$\begin{aligned}
p(y|\mu,\hat{\sigma}) &= \frac{1}{(2\pi)^{25}\cdot\sigma^{50}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{50}(y_i^2-2\mu y_i+\mu^2)} \\
p(\mu|\hat{\sigma}) &= \frac{1}{\sqrt{2\pi}\hat{\sigma}} e^{-\frac{1}{2}(\frac{\mu-11}{\hat{\sigma}})^2} \\
p(\hat{\sigma}) &= \frac{\hat{\sigma}^{8-1}(1-\hat{\sigma})^{2-1}}{\mathrm{B}(8,2)}, \quad \text{where} \quad \mathrm{B} = \frac{\Gamma(8)\Gamma(2)}{\Gamma(8+2)}
\end{aligned}$$

As we now work with a hierarchical model, writing the explicit expressions will be very messy. We therefore keep the general mathematical expressions.

- **Hyper prior:**
  $p(\hat{\sigma})$

- **Prior:**
  $\Pi(\mu, \hat{\sigma}) = p(\mu|\hat{\sigma}) \cdot p(\hat{\sigma})$

- **Likelihood**:
  $p(y|\mu, \hat{\sigma})$

- **Posterior**:

$$p(\mu, \hat{\sigma}|y) = \frac{p(y|\mu, \hat{\sigma})\Pi(\mu, \hat{\sigma})}{\int\int p(y|\mu, \hat{\sigma}\Pi(\mu, \hat{\sigma})d\mu d\hat{\sigma}} = \frac{p(y|\mu, \hat{\sigma})p(\mu|\hat{\sigma})p(\hat{\sigma})}{\int\int p(y|\mu, \hat{\sigma})p(\mu|\hat{\sigma})p(\hat{\sigma})d\mu d\hat{\sigma}}$$

- **Prior predictive:**
  $p(\tilde{y}) = \int\int p(\tilde{y}|\mu, \hat{\sigma})\Pi(\mu, \hat{\sigma})d\mu d\hat{\sigma} = \int\int p(\tilde{y}|\mu, \hat{\sigma})p(\mu|\hat{\sigma})p(\hat{\sigma})d\mu d\hat{\sigma}$

- **Posterior predictive:**
  $p(\tilde{y}|y) = \int\int p(\tilde{y}|\mu, \hat{\sigma})p(\mu, \hat{\sigma}|y)d\mu d\hat{\sigma}$

4 Experiment

To gain data, we made the our competitive Rubik's cube solver solve the cube 50 times while it was being timed. The scrambles were random-generated by a computer, giving each data a different scramble for each solve. Our member is allowed to look at it for 15 seconds (this is a rule that applies in official competitions) before he started to solve the cube. We used a computer to time the solves, where the competitor starts and stops the timer himself. The dataset obtained can be found in the zip file.

5 Model selection

In order to decide which is the model that best fits the data we are going to use the concept of **Bayes factor**. We are going to do a really simple explanation of what the Bayes Factor is:

- Consider two hypotheses H_0 and H_1 (corresponding to two possible models) and some data y.
- The Bayes factor is a ratio of the likelihood probability of these two hypotheses and quantifies the evidence of data y for H_0 vs H_1.
- Its formal definition is given by:

$$K = \frac{Pr(D|M_1)}{Pr(D|M_2)} = \frac{\int Pr(\theta_1|M_1)Pr(D|\theta_1, M_1)d\theta_1}{\int Pr(\theta_2|M_2)Pr(D|\theta_2, M_2)d\theta_2}$$

  Then, the Bayes factor (BF) is the ratio of the **marginal likelihoods**:

$$p(x|H_i) = \int \underbrace{p(x|\theta, H_i)}_{Likelihood} \underbrace{p(\theta|H_i)}_{Prior} d\theta$$

- **Interpretation** of Bayes Factor:

Bayes Factor	Strength of evidence
$< 1:1$	Negative (supports H_1)
1:1 to 3:1	Barely worth mentioning
3:1 to 20:1	Substantial
20:1 to 150:1	Strong
$> 150:1$	Very Strong

Table 1: BF Interpretation

In our particular case we are going to compare our 4 models in pairs.
As the expression of $p(x|H_i)$ of our models are quite large and the software R is not very efficient with this type of expressions (and the results are not reliable), we have done the calculations of each value with the help of $WolframAlpha$.
Here we have the results of $p(x|H_i)$ obtained with the 4 models:

- **Model 1:**

$$p(x|H_1) = \int_0^{\infty} \frac{1}{(2\pi)^{25} \cdot 1.5^{50}} e^{-\frac{1}{2 \cdot 1.5^2} \sum_{i=1}^{50} (y_i^2 - 2\mu y_i + \mu^2)} d\mu = 1.81743 \times 10^{-39}$$

- **Model 2:**

$$p(x|H_2) = \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{51}{2}} 0.8 \cdot 1.5^{50}} exp\left\{ \frac{-\mu^2 + 2 \cdot 11\mu - 11^2}{2 \cdot 0.8^2} - \sum_{i=1}^{50} \frac{y_i^2 - 2\mu y_i + \mu^2}{2 \cdot 1.5^2} \right\} d\mu = 6.83327 \times 10^{-40}$$

- **Model 3:**

$$p(x|H_3) = \int_9^{13} \frac{e^{-\frac{1}{2 \cdot 1.5^2} \sum_{i=1}^{50} (y_i^2 - 2\mu y_i + \mu^2)}}{(2\pi)^{25} 1.5^{50} 4} d\mu = 4.54358 \times 10^{-40}$$

- **Model 4:**

$$p(x|H_4) = \int_{-\infty}^{\infty} \int_0^1 \frac{e^{-\frac{1}{2 \cdot 1.5^2} \sum_{i=1}^{50} (y_i^2 - 2\mu y_i + \mu^2)} e^{-\frac{1}{2}\left(\frac{\mu-11}{\sigma}\right)^2} \sigma^7 (1-\sigma)}{\frac{(2\pi)^{25} 1.5^{50} \left(\sqrt{2\pi}\sigma\right) 7!}{9!}} d\sigma d\mu = 6.78512 \times 10^{-40}$$

So, if we compare these results by pairs, we obtain:

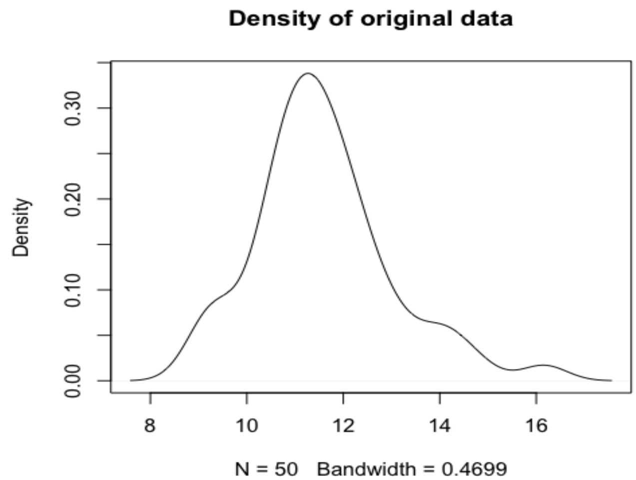
Models compared	**BF**
1 vs 2	2.659
1 vs 3	3.999
1 vs 4	2.678
2 vs 3	1.503
2 vs 4	1.007
3 vs 4	0.669

Table 2: Summary of model comparison

It seems like model 1 is the best model from Bayes factor, but in reality the value we obtain is not very representative. This is because model 1's prior is a non-informative prior, which will make it's contribution to Bayes factor much larger than the other models. The reason for this, is that the priors of the other models have an area of 1, while model 1's prior has an infinite area. We therefore choose to ignore that it was the best model according to Bayes factor, as Bayes factor cannot be applied to a model with a prior like model 1 (and still give correct results). We end opp choosing model 2, as this gives the second best result when using Bayes factor. Although it is barely preferred over the 4th model, we choose to not use the 4th model due to both its complexity with more levels and that model 2 had better results from using Bayes factor.

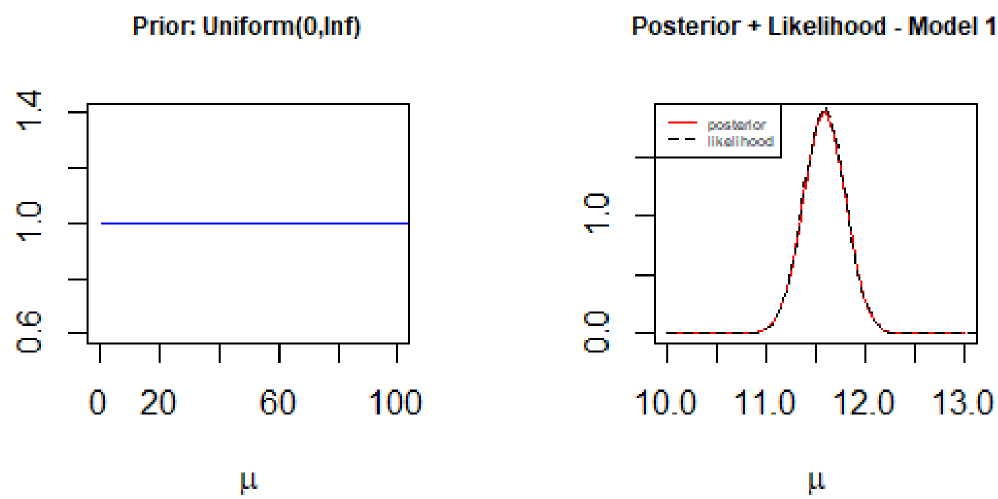
6 Results and Prediction

Let us first look at the density of the data:

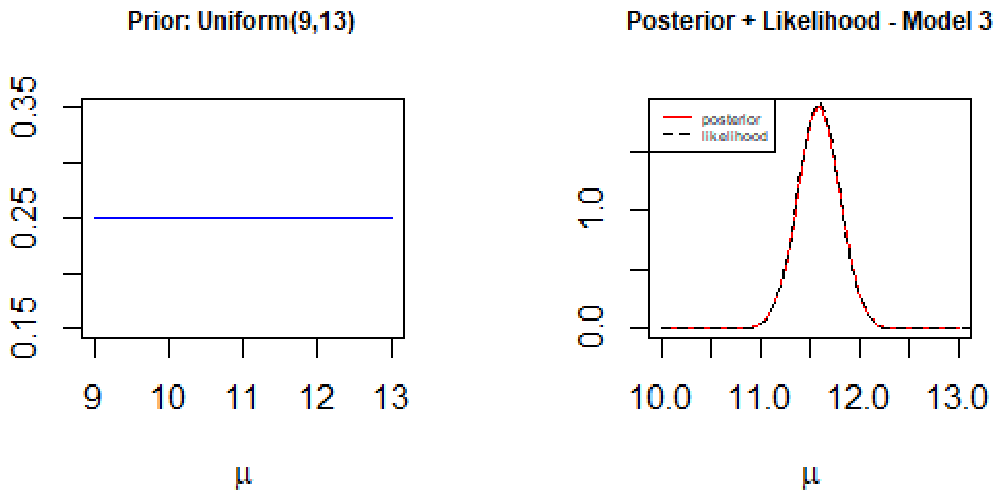


Before we look at our chosen model we (model 2) decided to plot some prior, likelihoods and posteriors for the other models:

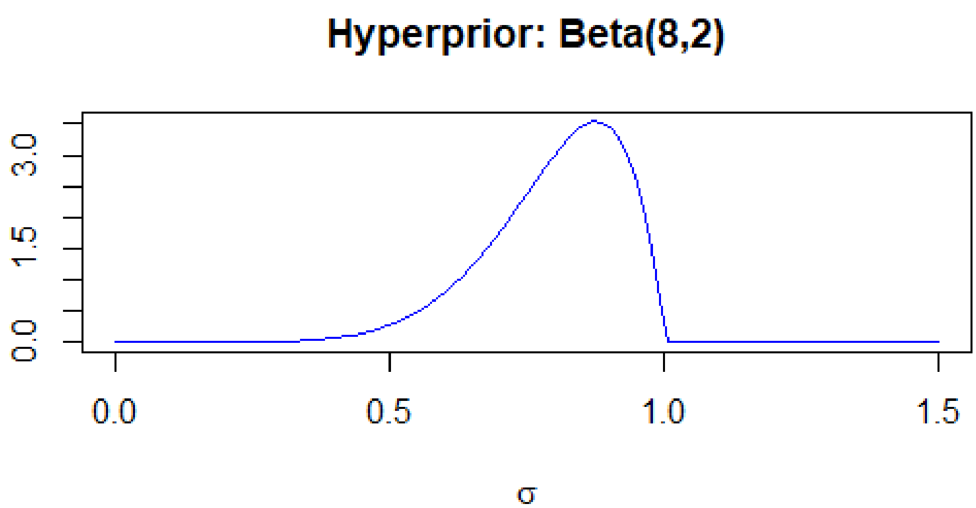
6.1 Model 1



6.2 Model 3

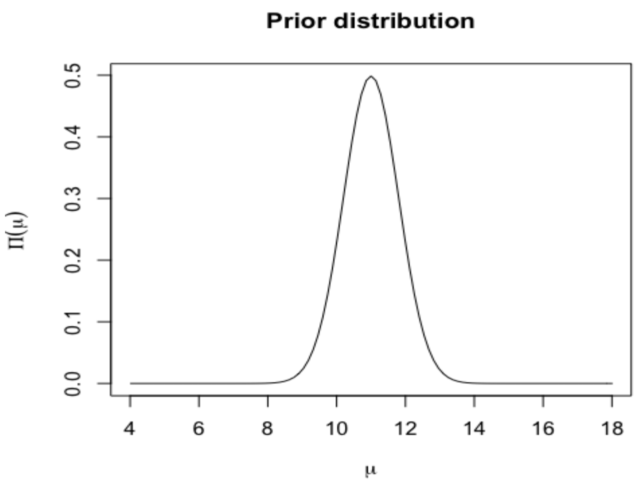


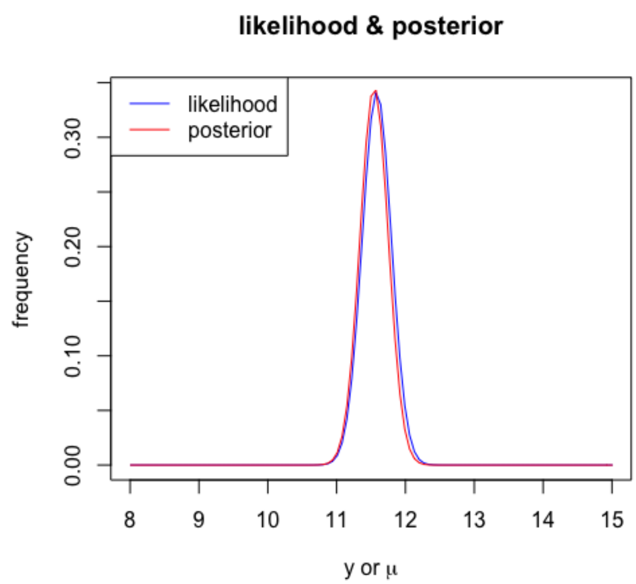
6.3 Model 4



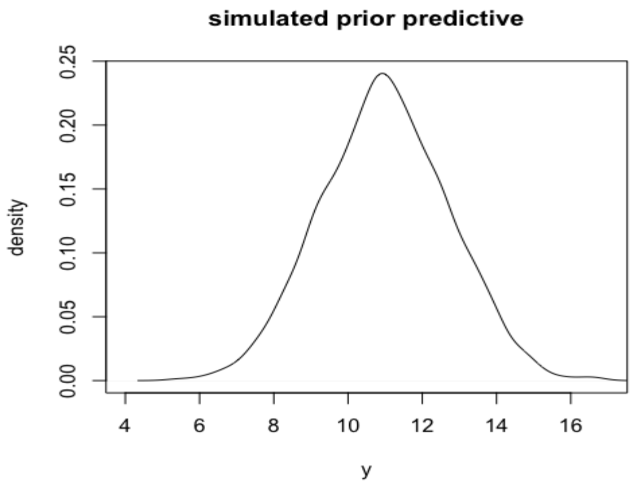
6.4 Model 2

As we chose model 2, lets look at some plots. The prior, likelihood and posterior is plotted below.

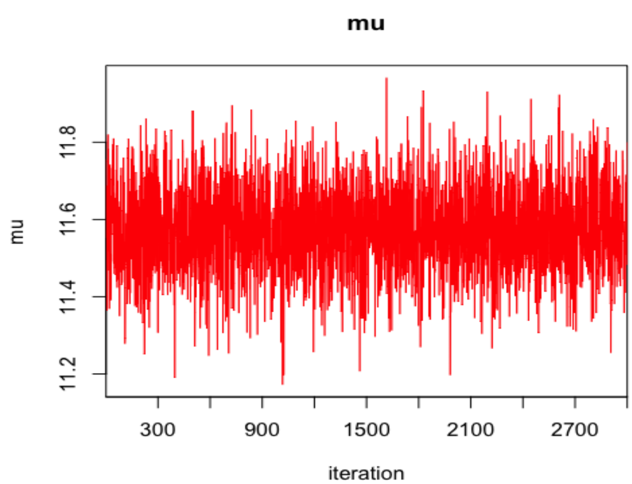




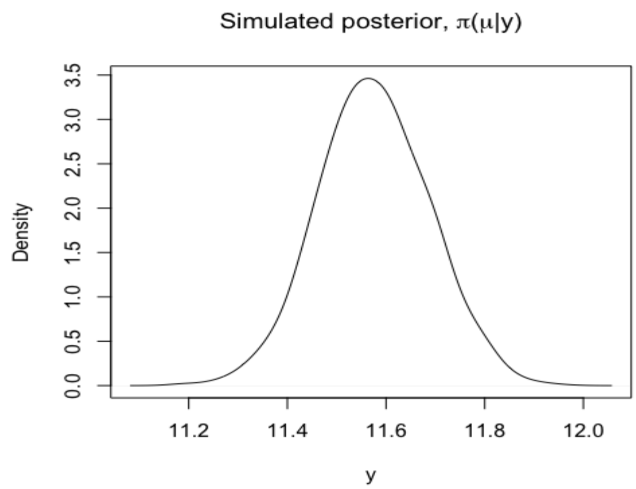
We then simulate 3000 μ-values from the normal distribution $\mathrm{N}(11, 0.8)$ (prior), and simulate y's from $\mathrm{N}(\mu, 1.5)$, where μ are the 3000 simulated values. The 3000 values for y are plotted as a density plot:

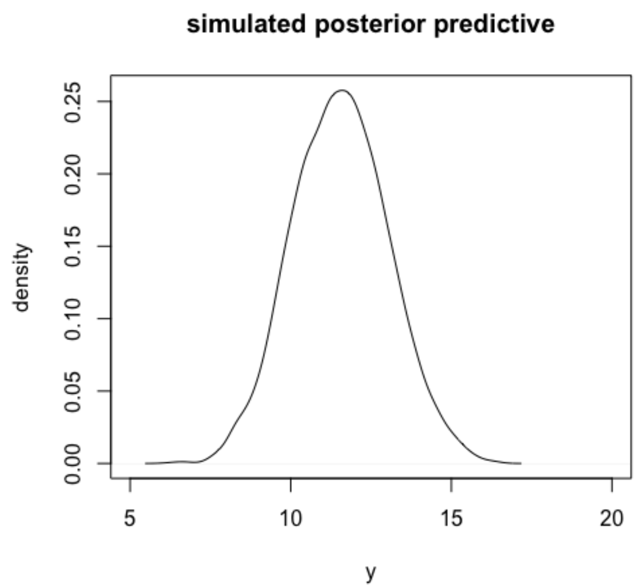


We now simulate μ with Markov chain Monte Carlo. This is done by iterating 3000 times. The result can be seen in the following plot:



Using the mcmc-simulated μ-values, we can now plot the posterior and posterior predictive. These are shown in the following plot:





7 Discussion

It seems like the simulated posterior predictive is much more narrow than the simulated prior predictive. Although the prior predictive seems to be centered approximately around the mean, it seems a bit too wide compared to the dataset, so we can see that the posterior predictive is a better predictive model. Before the project, we had the assumption that the solving times would be normally distributed with a mean that was approximately 11. This is why we constructed all of our models so that the expected value of the mean would be 11. The result we obtained had a mean of 11.5834, which made the likelihood, posterior and predictive posterior shift to the right. The prior and the predictive prior still has a mean around 11.

8 Conclusion

Our goal for the project was to predict how the future Rubik's cube solving times was distributed, and the posterior predictive seems like a normal distribution. This is very interesting, as finding out how a competitor's result is distributed, might help us to find out how he can improve. For example if he has a very big standard deviation, it might be interesting to see how he can improve his worst solves. If he has a narrow standard deviation, maybe he can look at how to improve his mean.

9 Code

The code can be found in the zip file.