

# CHL5223 - Applied Bayesian Methods - Assignment 1

Luis Correia - Student No. 1006508566

January 18th 2020

## Question 1

Preamble: You have been hired to predict the results of an election. The election is for town select- person. There are three select-persons who make up the the town council and who run the town. There is a division between the purple party and the brown party. If one side has a majority, they will control the issues in the town. Each select-person is elected in one of three town districts and both parties have candidates running in each district. If a candidate wins a majority of votes in a district, that candidate will be the select-person for that district. Also, assume that there will be 5001 citizens voting in each district.

The main observed random variable is  $X_i$ . This is the number of votes cast for the purple party in the  $i_{th}$  district. Let the probability that someone from district  $i$  votes for purple be equal to  $\theta_i$  where  $\theta_i$  has an unknown distribution.(That is, you can assume that the distribution is a beta distribution but you don't know the parameters of the distribution and you would like to learn about  $\theta_i$ .) Given the value of  $x_i$  one can define the value  $D_i$  as one if the purple party is elected and 0 otherwise. Also, one can define the value  $T$  as one if the purple party is elected and 0 otherwise.

### item (a)

Write down the probability generating model for the above scenario. It should include something about the conditional distribution of  $X_i$  and the distribution of  $\theta_i$ . You can also add in your model that  $D_i$  and  $T$  are deterministic (non-random) functions of the random variables  $X_i$ .

{Solution.}

In order to define the probability generating model we will consider the following definitions:

- $X_i$ : # of votes cast for Purple party in  $i^{th}$  district, with  $i \in \{1, 2, 3\}$ ;
- $D_i$ : equals 1 if Purple party is elected in  $i^{th}$  district, with  $i \in \{1, 2, 3\}$ ;
- $T$ : equals 1 if Purple *is elected*, which I interpreted as *the Purple party controls at least 2 districts*;
- The total number of voters is  $n = 5001$  in each district;

I will also define a binary variable  $V_{k,i}$  representing the  $k^{th}$  vote in the  $i^{th}$  district, such that:

$$V_{k,i} = \mathbb{1}(k) = \begin{cases} 1, & \text{if the } k^{th} \text{ vote is for Purple party} \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

We will also consider the probability of  $V_{k,i} = 1$  is  $\theta_i$  which represents the probability of  $k^{th}$  vote is cast in favor of Purple party in the  $i^{th}$  district.

$$\implies P(V_{k,i} = 1) = \theta_i \text{ and } P(V_{k,i} = 0) = (1 - \theta_i), k = 1, 2, \dots, n \text{ and } i = 1, 2, 3 \quad (2)$$

Using (1) and (2) above, we have:

$$X_i = \sum_{k=1}^n V_{k,i} \quad (3)$$

From (3) we can conclude the **likelihood distribution** of our model is given by:

$$p(X_i|\theta_i) \sim \text{Binomial}(X_i|\theta_i, n), \text{ with } i = 1, 2, 3 \text{ and } n = 5001 \quad (4)$$

We know that the conjugate prior of  $\theta_i$  has a Beta distribution with hyper-parameters  $\alpha_i$  and  $\beta_i$ , which implies the **prior distribution** of  $\theta_i$  is:

$$p(\theta_i|\alpha_i, \beta_i) \sim \text{Beta}(\theta_i|\alpha_i, \beta_i), \text{ with } i = 1, 2, 3 \quad (5)$$

and the **posterior distribution** is given by:

$$p(\theta_i|x_i) \sim \text{Beta}(\theta_i|\alpha_i + x_i, \beta_i + n - x_i), \text{ with } i = 1, 2, 3 \text{ and } n = 5001 \quad (6)$$

The deterministic variables  $D_i$  and  $T$  which also compound our model can be expressed as:

$$D_i|x_i = \mathbb{1}(x_i > 2500) = \begin{cases} 1, & \text{if } x_i > 2500 \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

and

$$T|D_1, D_2, D_3 = \mathbb{1}\left(\sum_{i=1}^3 D_i \geq 2\right) = \begin{cases} 1, & \text{if } \sum_{i=1}^3 D_i \geq 2 \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

### item (b)

Prior distributions. Create two different analyses. One based on a “non-informative prior” and the other on an informative prior. Justify your answer. For the informative prior consider the following information. In the past, the two parties have been quite balanced. In the past, each party would get between 40% to 60% of the votes. Use this information to create an informative prior. (*To receive full credit for your specification of your informative prior, you need to present a prior which has approximately 95% of its mass between .40 and .60.*)

{Solution.}

I will divide the answer to this question in 02 parts:

- Case 1: Non-informative prior - In this case I will consider we don't have any useful information for this prior;

- Case 2: Informative prior - In this case, I will translate the assumptions proposed by the question into useful statistics language, as seen in Lecture 2:

1. *In the past, the two parties have been quite balanced.*
2. *In the past, each party would get between 40% and 60%.*

### Case 1: Non-informative prior

For this case we will consider the prior distribution of  $\theta_i$  as an *Uniform* as seen in Lecture 1. Another justification for that approach was provided in Laplace's argument of "if nothing is known about  $\theta$ , the uniform specification is appropriate" (this is also called the "principle of insufficient reason")<sup>[1]</sup>

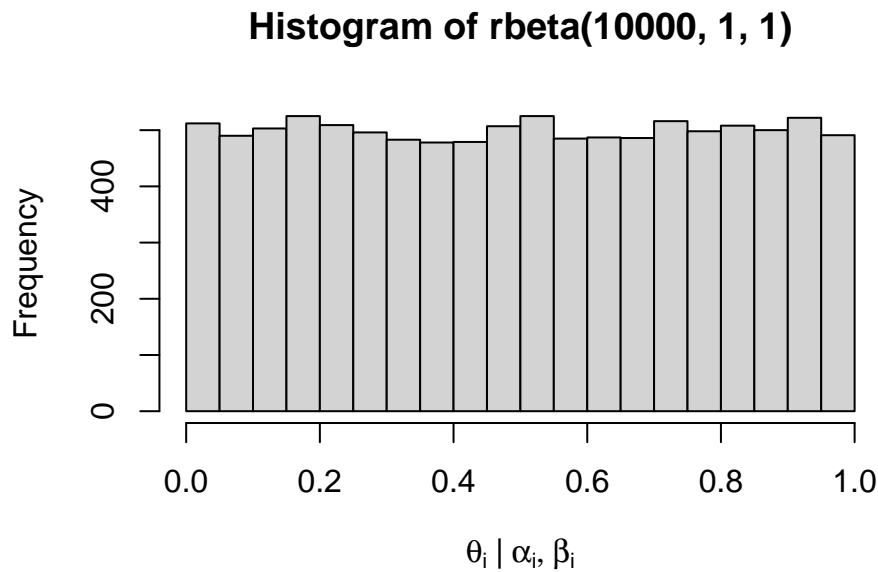
With this in mind, we have:

$$\alpha_i = \beta_i = 1, \text{ for } i = 1, 2, 3$$

and our **Non-Informative prior** can be written as follows:

$$\Rightarrow p(\theta_i | \alpha_i, \beta_i) \sim \text{Beta}(\theta_i | 1, 1), \text{ for } i = 1, 2, 3 \quad (9)$$

Graphically, by simulating  $N = 10,000$  samples of  $\theta_i$  we can see it follows a standard Uniform distribution, as expected.



### Case 2: Informative prior

In this case, we will translate the statements as follows:

1. *In the past, the two parties have been quite balanced.* - This may be considered as the expectation of  $\theta_i$  are equal for Purple and Brown parties. In other words, we can assume:

$$\mathbb{E}(\theta_i) = 0.5$$

Remember that  $\theta_i$  is distributed according  $Beta(\theta_i|\alpha_i, \beta_i)$ , then follows:

$$\begin{aligned}\frac{\alpha_i}{\alpha_i + \beta_i} &= 0.5 \\ \alpha_i &= 0.5 \times (\alpha_i + \beta_i) \\ 0.5 \times \alpha_i &= 0.5 \times \beta_i \\ \implies \alpha_i &= \beta_i, \text{ for } i = 1, 2, 3\end{aligned}\tag{10}$$

2. In the past, each party would get between 40% and 60%. - in this case, is reasonable to assume that  $\theta_i$  lies in the interval  $[0.4; 0.6]$ . This statement can also be re-written as:

$$\mathbb{P}(\theta_i \in [0.4; 0.6]) = 0.95$$

which is equivalent to:

$$\mathbb{P}(\theta_i \leq 0.6) - \mathbb{P}(\theta_i \leq 0.4) = 0.95, \text{ for } i = 1, 2, 3$$

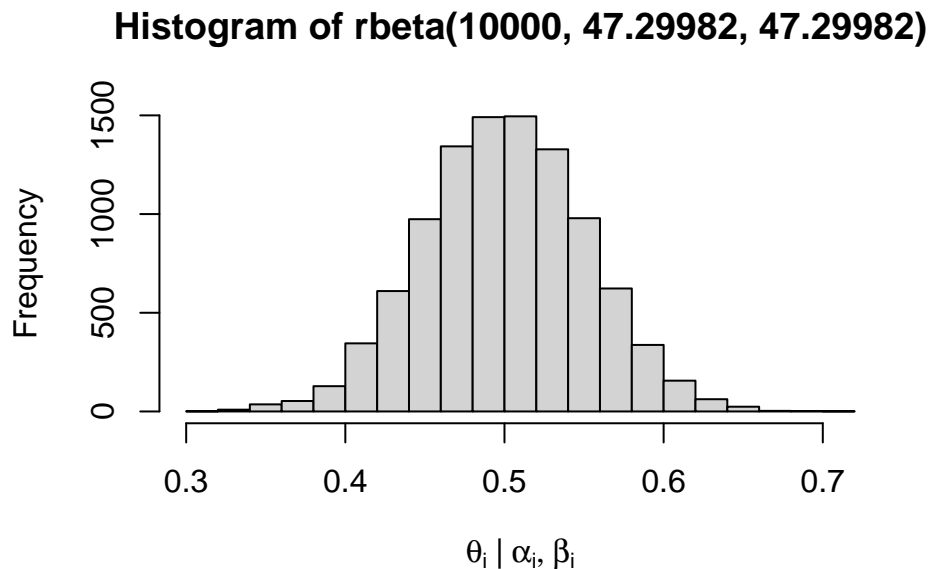
Using the *Bisection Algorithm*<sup>1</sup> seen in class and the information in statements #1 and #2, we obtained:

```
## The prior distribution of Theta is Beta( 47.29982 , 47.29982 )
## This distribution guarantees that 95% of its mass is between 0.40 and 0.60 as we can see below:
## -> pbeta(0.4, 47.29982, 47.29982)= 0.025 and pbeta(0.6, 47.29982, 47.29982)= 0.975
## -> qbeta(c(0.025,0.975), 47.29982, 47.29982)= 0.4 0.6
```

This implies that our **Informative Prior** can be written as follows:

$$\implies p(\theta_i|\alpha_i, \beta_i) \sim Beta(\theta_i|47.29982, 47.29982), \text{ for } i = 1, 2, 3\tag{11}$$

By simulating  $N = 10,000$  from the prior obtained we can see graphically the desired distribution.



<sup>1</sup>Please refer to Appendix for the R-Code used.

**item (c)**

Posterior distribution for the voting percentage for each district. For each district, a simple random sample of the citizen of each district is asked whom they plan to vote for. In district one: 53 said purple and 45 said brown; in district two: 72 said purple and 78 said brown; and in district three: 18 said purple and 22 said brown. Calculate the posterior probability for the percent who will vote for purple in each district. In reporting your results, provide the distribution of the posterior and the value of the parameters and also report appropriate statistics for these distribution (which includes the posterior mean, standard deviation, and some kind of 95% interval). (Note: do this for each of the two priors specified in the first part.)

{Solution.}

The sample from citizens of each district has shown the following results, which will be used to calculate the respective posterior distribution.

Table 1: Sampling of Vote Intention by District

District	Purple	Brown	Total
1	53	45	98
2	72	78	150
3	18	22	40

Additionally, we included for each district the posterior probability that Purple party wins on each district. This probability, as seen in “Supplemental Lecture”, follows a *Beta-Binomial* distribution, and it is more spread because it incorporates the uncertainty of each  $\theta_i$ .

**Case 1: Non-Informative prior**

Using the prior obtained in (9), we calculated the **posterior distribution for each district**.

For **District #1** we have:

$$\implies p(\theta_1 | \alpha_1, \beta_1, x_1) \sim \text{Beta}(\theta_1 | 54, 46) \quad (12)$$

... and the summary statistics are:

```
##
##
## --- Summary for District 1 ---
##
## Posterior Probability for the percent o voters for Purple party in Dist.#1 is 0.540
##
## Mean= 0.540
##
## Std. Deviation= 0.0495924
##
## 95% Credible Region ( 0.4423 , 0.6362 )
```

```
##
## Posterior Probability Purple party win in Dist.#1 is 0.786962
```

For **District #2** we have:

$$\implies p(\theta_2|\alpha_2, \beta_2, x_2) \sim \text{Beta}(\theta_2|73, 79) \quad (13)$$

... and the summary statistics are:

```
##
##
## --- Summary for District 2 ---

##
## Posterior Probability for the percent o voters for Purple party in Dist.#2 is 0.480263

##
## Mean= 0.480263

##
## Std. Deviation= 0.0403911

##
## 95% Credible Region ( 0.4015 , 0.5596 )

##
## Posterior Probability Purple party win in Dist.#2 is 0.315313
```

For **District #3** we have:

$$\implies p(\theta_3|\alpha_3, \beta_3, x_3) \sim \text{Beta}(\theta_3|19, 23) \quad (14)$$

... and the summary statistics are:

```
##
##
## --- Summary for District 3 ---

##
## Posterior Probability for the percent o voters for Purple party in Dist.#3 is 0.452381

##
## Mean= 0.452381

##
## Std. Deviation= 0.0759027

##
## 95% Credible Region ( 0.3066 , 0.6025 )

##
## Posterior Probability Purple party win in Dist.#3 is 0.267195
```

## Case 2: Informative prior

Using the prior obtained in (11), we calculated the **posterior distribution for each district**.

For **District #1** we have:

$$\implies p(\theta_1|\alpha_1, \beta_1, x_1) \sim \text{Beta}(\theta_1|100.2998, 92.29982) \quad (15)$$

... and the summary statistics are:

```
##
##
## --- Summary for District 1 ---
##
## Posterior Probability for the percent o voters for Purple party in Dist.#1 is 0.520768
##
## Mean= 0.520768
##
## Std. Deviation= 0.035904
##
## 95% Credible Region ( 0.4503 , 0.5909 )
##
## Posterior Probability Purple party win in Dist.#1 is 0.714646
```

For **District #2** we have:

$$\implies p(\theta_2|\alpha_2, \beta_2, x_2) \sim \text{Beta}(\theta_2|119.2998, 125.2998) \quad (16)$$

... and the summary statistics are:

```
##
##
## --- Summary for District 2 ---
##
## Posterior Probability for the percent o voters for Purple party in Dist.#2 is 0.487735
##
## Mean= 0.487735
##
## Std. Deviation= 0.0318952
##
## 95% Credible Region ( 0.4254 , 0.5503 )
##
## Posterior Probability Purple party win in Dist.#2 is 0.353729
```

For **District #3** we have:

$$\implies p(\theta_3|\alpha_3, \beta_3, x_3) \sim \text{Beta}(\theta_3|65.29982, 69.29982) \quad (17)$$

... and the summary statistics are:

```
##
##
## --- Summary for District 3 ---

##
## Posterior Probability for the percent o voters for Purple party in Dist.#3 is 0.485141

##
## Mean= 0.485141

##
## Std. Deviation= 0.0429189

##
## 95% Credible Region ( 0.4014 , 0.5693 )

##
## Posterior Probability Purple party win in Dist.#3 is 0.366426
```

#### item (d)

Given the above information, provide the probability that the purple party will have a majority in the town council. Also, provide the probability that purple will win each of the districts. Do this by simulation. That is, simulate 10,000 elections. For each simulated election, generate a probability,  $\theta_i$ , which is the probability of a citizen voting for the purple party in district  $i$ . Then, generate the number of citizens voting for purple in each district. From there, one can elect either the purple candidate or the brown candidate for each district for each simulated election. Finally, one can see which party had a majority in the town council in the simulated election. (Note: do this for each of the two priors specified in the first part.) (*For full credit on this simulation, you need to show that for each simulated election you sample a  $\theta_i$  and the number of citizens who vote in each district. Also, no, I don't want to see the actual list of 10,000 sampled values.*)

{Solution.}

For this question we will use the posterior distributions obtained in cases #1 and #2, i.e., Non-Informative and Informative Priors, respectively to simulate  $N = 10,000$  sample elections. After doing this, we will use the simulated elections to calculate the proportion of 'wins' of Purple party on each district and calculate the proportion of times it rules the town.

#### Case 1: Non-Informative prior

For this case, we will use the posterior distributions obtained in (12), (13) and (14) for each district.

The algorithm<sup>2</sup> may be summarized as follows:

- Loop:  $n = 1$  to 10000
  - Sample  $\hat{\theta}_i$  (=probability to vote Purple) for each district from the posteriors obtained in item (c)
    - \* District-1: `rbeta(1,54,46)`
    - \* District-2: `rbeta(1,73,79)`
    - \* District-3: `rbeta(1,19,23)`

---

<sup>2</sup>The respective R-code can be seen in Appendix.



- for each district, use the  $\hat{\theta}_i$  from the posterior to sample voters for each District with :
  - \*  $\hat{\theta}_i$  to 'vote for Purple party';
  - \*  $(1 - \hat{\theta}_i)$  otherwise
  - \* R-Code: `rbinom(1,size=5001,prob= $\theta_i$ )`
- Summarize the result of the election by summing how many '1' were obtained for each district
- Store the current result of the election of each district
- End Loop
- Calculate the proportion of 'wins' for Purple party
- Estimate the probability of Purple party win in each district  $\hat{\phi}_i = \frac{\text{No. of Victories of Purple Party}}{10,000}$
- Calculate the probability of Purple party wins in at least 2 district
- Report the results

The probability that Purple party have the majority in the town is calculated as follows:

$$\mathbb{P}(\text{Purple party controls Town Council}) = \mathbb{P}(\text{Purple party wins in at least 02 districts})$$

$$\mathbb{P}(\text{Purple party wins in at least 02 districts}) = \mathbb{P}\left(\sum_{i=1}^3 D_i \geq 2\right)$$

... this is equivalent to:

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^3 D_i \geq 2\right) &= \mathbb{P}(D_1 = 1, D_2 = 1, D_3 = 0) + \\ &\quad \mathbb{P}(D_1 = 1, D_2 = 0, D_3 = 1) + \mathbb{P}(D_1 = 0, D_2 = 1, D_3 = 1) + \\ &\quad \mathbb{P}(D_1 = 1, D_2 = 1, D_3 = 1) \end{aligned}$$

$$\begin{aligned} \Rightarrow \mathbb{P}\left(\sum_{i=1}^3 D_i \geq 2\right) &= \mathbb{P}(X_1 > 2500, X_2 > 2500, X_3 \leq 2500 | \theta_1, \theta_2, \theta_3) + \\ &\quad \mathbb{P}(X_1 > 2500, X_2 \leq 2500, X_3 > 2500 | \theta_1, \theta_2, \theta_3) + \\ &\quad \mathbb{P}(X_1 \leq 2500, X_2 > 2500, X_3 > 2500 | \theta_1, \theta_2, \theta_3) + \\ &\quad \mathbb{P}(X_1 > 2500, X_2 > 2500, X_3 > 2500 | \theta_1, \theta_2, \theta_3) \end{aligned}$$

$$\begin{aligned} \Rightarrow \mathbb{P}(\text{Purple party controls Town Council}) &= \mathbb{P}(X_1 > 2500, X_2 > 2500, X_3 \leq 2500 | \theta_1, \theta_2, \theta_3) + \\ &\quad \mathbb{P}(X_1 > 2500, X_2 \leq 2500, X_3 > 2500 | \theta_1, \theta_2, \theta_3) + \\ &\quad \mathbb{P}(X_1 \leq 2500, X_2 > 2500, X_3 > 2500 | \theta_1, \theta_2, \theta_3) + \\ &\quad \mathbb{P}(X_1 > 2500, X_2 > 2500, X_3 > 2500 | \theta_1, \theta_2, \theta_3) \end{aligned} \quad (18)$$

By using the proportion of wins on each simulation we can estimate the probabilities requested obtaining the following results:

The frequencies observed of Purple party on each district is as follows:

Table 2: Purple Party wins in 10,000 simulations

District	Freq	Prop
1	7806	0.7806
2	3226	0.3226
3	2694	0.2694

The frequencies observed of Purple party won 02 districts is as follows:

Table 3: Purple Party wins at least 02 districts in 10,000 simulations

District	Freq	Prop
1 & 2	2505	0.2505
1 & 3	2087	0.2087
2 & 3	866	0.0866

Table 4: Purple Party wins exactly 02 districts / All districts in 10,000 simulations

District	Freq	Prop
Only 1 & 2	1823	0.1823
Only 1 & 3	1405	0.1405
Only 2 & 3	184	0.0184
All	682	0.0682

Using the proportions from the simulation and (18), the probabilities requested are as follows:

```
##
## Probability of Purple Win District #1= 0.7806

##
## Probability of Purple Win District #2= 0.3226

##
## Probability of Purple Win District #3= 0.2694

##
##
## Probability that the Purple Party have majority in Town Council= 0.4094
```

## Case 2: Informative prior

For this case, we will use the posterior distributions obtained in (15), (16) and (17) for each district.

The algorithm is the same as in Case 1 (above).

The frequencies observed of Purple party on each district is as follows:

Table 5: Purple Party wins in 10,000 simulations

District	Freq	Prop
1	7146	0.7146
2	3512	0.3512
3	3540	0.3540

The frequencies observed of Purple party won 02 districts is as follows:

Table 6: Purple Party wins at least 02 districts in 10,000 simulations

District	Freq	Prop
1 & 2	2491	0.2491
1 & 3	2538	0.2538
2 & 3	1238	0.1238

Table 7: Purple Party wins exactly 02 districts / All districts in 10,000 simulations

District	Freq	Prop
Only 1 & 2	1609	0.1609
Only 1 & 3	1656	0.1656
Only 2 & 3	356	0.0356
All	882	0.0882

Using the proportions from the simulation and (18), the probabilities requested are as follows:

```
##
## Probability of Purple Win District #1= 0.7146

##
## Probability of Purple Win District #2= 0.3512

##
## Probability of Purple Win District #3= 0.354

##
##
## Probability that the Purple Party have majority in Town Council= 0.4503
```

It is interesting to note that the probabilities obtained by simulation are pretty much aligned with the *Beta-Binomial* distribution of each posterior distribution.

Table 8: Simulated vs. and Beta-Binomial Probability (Case 1: Non-Informative Prior)

District	Simul	BetaBin
1	0.7806	0.7870
2	0.3226	0.3153
3	0.2694	0.2672

Table 9: Simulated vs. and Beta-Binomial Probability (Case 2: Informative Prior)

District	Simul	BetaBin
1	0.7146	0.7146
2	0.3512	0.3537
3	0.354	0.3664

By plotting the sample densities obtained in simulations, we can graphically see the difference from **Case 1 (Non-Informative Prior)** vs. **Case 2 (Informative Prior)**.

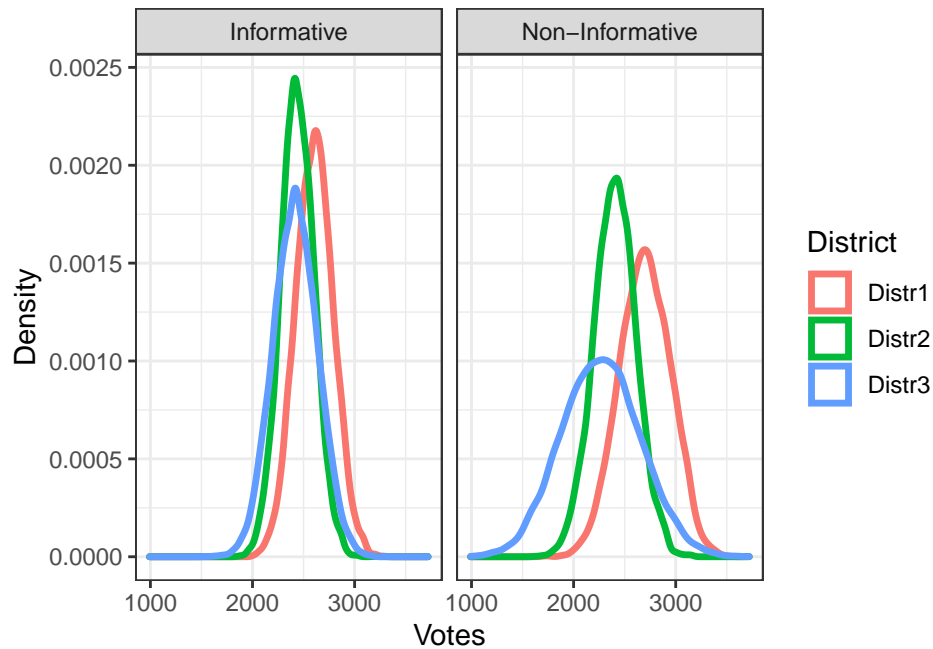


Figure 1: Sample Density for each District

The use of Non-Informative priors lead to “flatten” posterior densities when compared to Informative Priors. We can also see that for each district we have different beliefs for the expected number of votes for Purple Party in Non-Informative priors, in contrast with the other group, which presented “more stacked” posteriors with higher probability level as well.

Table 10: Comparison of Amplitude of 95% Credible Interval

District	NonInf	Inform
1	0.1940	0.1406
2	0.1581	0.1249
3	0.2959	0.1680

The amplitude of 95% credible Interval are smaller for *Informative Prior* when compared with *Non-Informative*, confirming the better accuracy of posterior distributions when using Informative priors.

## Question 2

(Some basic Bayesian questions) Let's calculate the posterior distribution of the average height (in inches) of a group of people. Suppose that we have two different data-sets. The first data-set is:  $D1 = (53, 49, 63, 72, 55, 65)$ , and the second is:  $D2 = (28, 27, 36, 42, 25, 35)$ . Consider the following three priors for this data and then answer the question below.

**Prior 1:** Assume the heights  $X_i$  of people from a particular population follows a normal distribution with  $\mu$  and precision  $\tau$ . Also, assume that the parameter  $\tau$  is known and that  $\mu$  has a normal prior distribution with mean  $\mu_0$  and precision  $\tau_0$ . Assume that it is believed that the *average* height is 66 inches (about 182cm). Also, assume that it is generally thought that the 95% of the  $X_i$ 's are between 54 and 78 inches. Then the posterior standard deviation is about one forth of that range and so the standard deviation of the  $X_i$ 's given  $\mu$  is about 6 inches, so the prior belief is that the fixed parameter  $\tau$  is about  $1/36$ . For the value of  $\mu_0$  it is believed that the true value  $\mu$  is between 63 inches and 69 inches (with about 95% probability). So, the value of  $\tau_0$  is about  $4/9$ .

**Prior 2:** Assume the heights,  $X_i$ , are distributed with mean  $\mu$  and precision  $\tau$ . Also, assume that  $\tau$  has gamma prior distribution with parameters (shape and rate) of  $\alpha$  and  $\beta$ , and  $\mu$  has a normal distribution given that  $\tau = t$  with mean  $\mu_0$  and precision  $\theta\tau$ . As with prior 1, we let  $\mu_0$  equal 66 inches. Also, from the reasoning in prior 1, we will assume that the prior mean of  $\tau$  is  $1/36$ . Since the mean of this gamma distribution is  $\alpha/\beta$ , then one possibility is to have  $\alpha = 1$  and  $\beta = 36$ . We can use R or **Spplus** to check the distribution of the standard deviation of  $X_i$  given  $\mu$  which we figured should be about 6. The following R/Spplus command can be used to check the definition:

```
quantile(1/sqrt(rgamma(10000,1,36)),probs=c(0,.025,.25,.50,.75,.975,1))
```

We see that there is very good support in the neighbor of 6 for the prior distribution. To finish this prior, we need a value for the parameter  $\theta$ . Now it could be argued that the values  $X_i$  given  $\mu$  should be much more spread out than the values of  $\mu$  around its mean. So, we will keep  $\theta$  greater than one. Here, let us use  $\theta$  to be about 4 and that leads to the following summary statistics for the prior distribution of the standard deviation of  $\mu$  given  $\mu_0$ :

```
quantile(1/sqrt(4*rgamma(10000,1,36)),probs=c(0,.025,.25,.50,.75,.975,1))
```

**Prior 3:** This prior uses the same basic model as prior as was used in Prior 2, but  $\mu_0$  equals 66,  $\theta = 0.1$ ,  $\alpha = .001$ , and  $\beta = .001$ .

### item (a)

For each of the three priors and for the two data sets (so there are 6 combinations in all), calculate the posterior mean, standard deviation, and a 95% credible region for average height. (Note: the standard precision of the t-distribution is not  $1/\text{variance}$ .) You may calculate these values by finding the exact values using analytic methods (algebra) or by performing simulations. (You can also do it both ways if you wish.).

{*Solution.*}

Our problem consists in estimate the average height of a certain population using Bayesian Methods.

Let  $X_i$ :Height of people from a particular population, be the variable of interest and for that we have 02 data-sets containing samples of height of this population:

- $D1 = (53, 49, 63, 72, 55, 65)$
- $D2 = (28, 27, 36, 42, 25, 35)$

In this case our **likelihood** distribution of  $\mu$  given  $\tau$  is:

$$p(X_i|\mu, \tau) \sim N\left(\mu, \frac{1}{\tau}\right) \quad (19)$$

We will consider 03 priors to analyze these data-sets.

### Prior #1

For Prior #1 we have the following assumptions:

In this prior we will consider  $\tau$  **known**, which implies:

$$X_i \sim N\left(\mu, \frac{1}{\tau}\right), \text{ with } \tau > 0 \text{ known}$$

### Summary of Beliefs

- Belief#1 - The average height is 66 inches ( $\sim 182\text{cm}$ )
- Belief#2 - The 95% of  $X_i$ 's lies between 54 and 78 inches.

Using the *Summary of Beliefs* we can presume the following prior distributions:

$$p(\mu|\mu_0, \tau_0) \sim N\left(\mu_0, \frac{1}{\tau_0}\right) \quad (20)$$

and the following prior information:

$$\mu_0 = 66, \tau = \frac{1}{36} \text{ and } \tau_0 = \frac{4}{9} \quad (21)$$

With this information, we conclude that the **posterior distribution of  $\mu$**  given the  $\mathbf{X}^{(D_i)}$  with  $i = 1, 2$  is:

$$p(\mu|\mathbf{X} = \mathbf{x}^{(D_i)}, n, \tau, \mu_0, \tau_0) \sim N\left(\frac{\tau_0\mu_0 + n\tau\bar{\mathbf{x}}^{(D_i)}}{\tau_0 + n\tau}, \frac{1}{\tau_0 + n\tau}\right), \text{ with } i = 1, 2 \quad (22)$$

Using (21) and (22) for both databases we obtained the following:

```
## -----
##
## Summary Statistics for Data-Set1- ( 53 49 63 72 55 65 )
##
## Posterior Mean is  64.2273
##
## Posterior Std. Deviation is  1.2792
##
## 95% Credible Region for mean ( 61.720 , 66.734 )
##
## -----
```

```
##
## Summary Statistics for Data-Set2- ( 28 27 36 42 25 35 )

##
## Posterior Mean is 56.7727

##
## Posterior Std. Deviation is 1.2792

##
## 95% Credible Region for mean ( 54.266 , 59.280 )
```

### Prior #2

For this prior, we have both  $\mu$  and  $\tau$  unknown, i.e.:

$$X_i \sim N\left(\mu, \frac{1}{\tau}\right), \text{ with } -\infty < \mu < \infty \text{ and } \tau > 0 \text{ unknown.}$$

... which will lead us to the following prior distributions:

$$p(\tau|\alpha, \beta) \sim \text{Gamma}(\alpha, \beta) \quad (23)$$

and

$$p(\mu|\tau = t, \theta, \mu_0, \tau_0) \sim N\left(\mu_0, \frac{1}{\theta t}\right) \quad (24)$$

In order to preserve the comparability of the models, we will inherit some assumptions from Prior#1 which are summarized as follows:

$$\mu_0 = 66, \theta = 4, \alpha = 1 \text{ and } \beta = 36 \quad (25)$$

With this information, we can calculate the **joint posterior distribution of  $\mu$**  given the  $\mathbf{X}^{(D_i)}$  with  $i = 1, 2$  which is compounded by the **marginal posterior distribution of  $\mu$**  given  $\tau = t$

$$p(\mu|\mathbf{X} = \mathbf{x}^{(D_i)}, n, \tau = t, \theta, \mu_0) \sim N\left(\frac{\theta\mu_0 + n\bar{\mathbf{x}}^{(D_i)}}{\theta + n}, \frac{1}{(\theta + n)t}\right), \text{ with } i = 1, 2 \quad (26)$$

... multiplied by the **posterior distribution of  $\tau$** , i.e.

$$p(\tau|\mathbf{X} = \mathbf{x}^{(D_i)}, n, \theta, \mu_0) \sim \text{Gamma}\left(\alpha + \frac{n}{2}, \beta + \frac{1}{2} \sum_{j=1}^n (x_j^{(D_i)} - \bar{\mathbf{x}}^{(D_i)})^2 + \frac{\theta n (\bar{\mathbf{x}}^{(D_i)} - \mu_0)^2}{2(\theta + n)}\right) \quad (27)$$

, with  $i = 1, 2$

Using (26) and (27) for both databases we obtained the following:

```
## -----
##
## Summary Statistics for Data-Set1- ( 53 49 63 72 55 65 )
```



```
##
## Posterior Mean is 62.100

##
## (Analytically)Posterior Std. Deviation is 3.01358

##
## (Simulated)Posterior Std. Deviation is 3.02083

##
## 95% Credible Region for mean ( 56.179 , 68.021 )

##
## 95% Credible Region for Std.Deviation ( 1.760 , 4.977 )

##
## -----

##
## Summary Statistics for Data-Set2- ( 28 27 36 42 25 35 )

##
## Posterior Mean is 45.700

##
## (Analytically) Posterior Std. Deviation is 7.11114

##
## (Simulated) Posterior Std. Deviation is 7.11871

##
## 95% Credible Region for mean ( 31.748 , 59.652 )

##
## 95% Credible Region for Std.Deviation ( 4.167 , 11.839 )
```

### Prior #3

For this prior, we will use the same approach in (26) and (27) and the following changes in the hyper-parameters:

$$\mu_0 = 66, \theta = 0.1, \alpha = 0.001 \text{ and } \beta = 0.001 \quad (28)$$

The results obtained are as follows:

```
## -----

##
## Summary Statistics for Data-Set1- ( 53 49 63 72 55 65 )

##
## Posterior Mean is 59.6066

##
## (Analytically) Posterior Std. Deviation is 3.92277
```

```

##
## (Simulated) Posterior Std. Deviation is  3.91637

##
## 95% Credible Region for mean ( 51.931 , 67.283 )

##
## 95% Credible Region for Std.Deviation ( 2.065 , 7.069 )

##
##
## -----

##
## Summary Statistics for Data-Set2- ( 28 27 36 42 25 35 )

##
## Posterior Mean is  32.7213

##
## (Analytically)Posterior Std. Deviation is  3.66231

##
## (Simulated)Posterior Std. Deviation is  3.64814

##
## 95% Credible Region for mean ( 25.571 , 39.872 )

##
## 95% Credible Region for Std.Deviation ( 1.927 , 6.547 )

```

### item (b)

For the six different posterior densities from the previous question, give the density plots and put all six on the same plot. (So, I want just one plot). Those densities can be either from the true, analytic distribution or they can be estimated from samples from the posterior distributions. For clarity, use a method to distinguish which is which. For example, you might consider the same color for the densities from the same prior and a different line type for the two different priors.

*{Solution.}*

For this question we will plot the densities of samples from the posterior distributions simulated with  $N = 10,000$ .

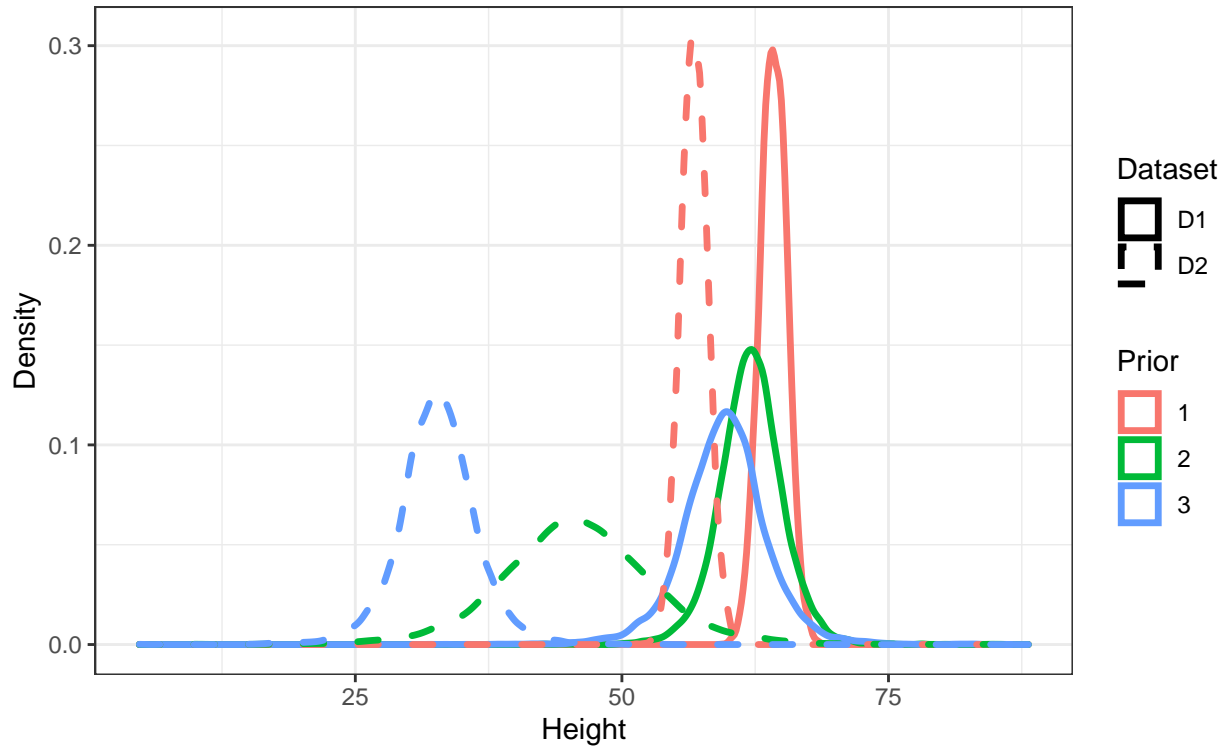


Figure 2: Sample Densities for each Data-Set, for each Prior

**item (c)**

Use data D1 and prior 3, find the predictive distribution for new observations. That is, using the posterior distribution, get the distribution of a new person from this population. To describe this predictive distribution, provide the mean, standard deviation, a 95% credible region and a density plot. These values can be obtained from either analytic methods, from sampling or a combination of analytic and sampling techniques. Don't forget to provide a description as to how you did this. (Either the formula or the sampling algorithm.)

{Solution.}

In this question it was calculated *analytically* and *by simulation* in order to compare the distributions.

Considering that the *predictive distribution* of new observations is given by the *marginal distribution* of  $\mu$ , as seen in supplemental paper from Lecture2, we need to calculate the expected value of  $\mu$  - which is a **t-Student** with non-centrality parameter given by:

$$p(\mu | \mathbf{X}_{D_1}, \mu'_0, \alpha', \beta', \theta') \sim \text{Std Student's-t}(2\alpha') \times \sqrt{\frac{\beta'}{\theta' \alpha'}} + \mu'_0$$

where

$$\mu'_0 = \frac{\theta \mu_0 + n \bar{x}_{D_1}}{\theta + n}$$

... degrees-of-freedom given by:

$$df = 2\alpha' = 2\alpha + n$$

... and with variance given by:

$$\mathbb{V}\text{ar}(\mu) = \left( \frac{2\alpha'}{2\alpha' - 2} \right) \left( \frac{\beta'}{\theta'\alpha'} \right)$$

For the sampling algorithm, we simulated  $N = 10,000$  samples from

$$\text{Std Student's-t}(6.002) \times 3.203194 + 59.60656$$

... obtaining the following results:

```
## -----
##
## Summary Statistics of Predictive Distribution
##
## #>>> Analytically ---
##
## Posterior Mean is 59.6066
##
## Posterior Std. Deviation is 3.92277
##
## 95% Credible Region for New Observation ( 51.769 , 67.444 )
##
##
## #>>> Simulated ---
##
## Posterior Mean is 59.6189
##
## Posterior Std. Deviation is 3.97992
##
## 95% Credible Region for New Observation ( 51.553 , 67.630 )
```

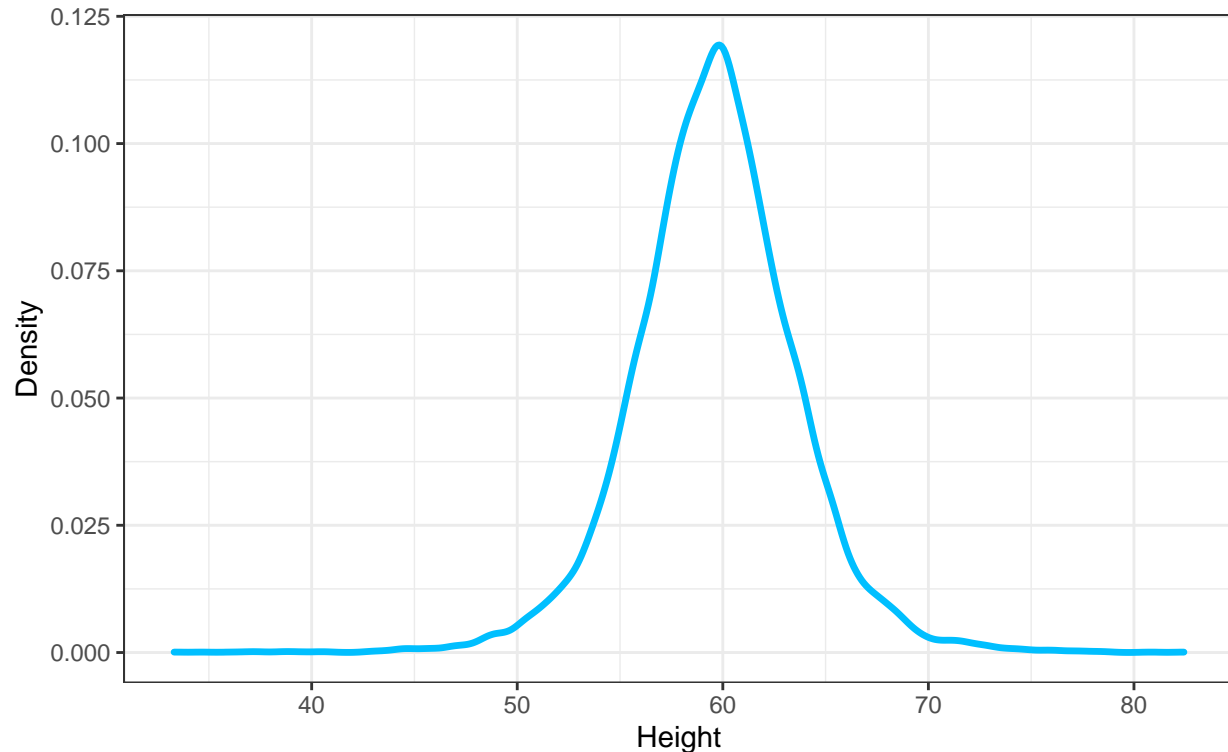


Figure 3: Predictive Distribution for Prior#3 and Data-set 1

**item (d)**

Please comment on the differences in the results obtained by using the different priors. To be more specific, try and imagine what kind of personal beliefs each of the three priors represent. So, each of the priors assumes that the “a prior” average height is 66 inches and then we are imagining that the person sees two different types of data. So, is there a person who might be describes as “not believing” the data when they see one set of data. So, perhaps that person might revise their estimate of the average height and perhaps have 95% credible region of the average which might not even contain any of the observed values. Also, note how their prior beliefs are reflected in the size of the posterior credible regions after seeing either of the two data sets. So, to answer this question, describe the beliefs of the three different people who would have each of the three different priors and state how they react to seeing each of the two data sets. (So, perhaps they were surprised by the data, they might have been suspicious of the data, or they might have seen the data as “confirming” their belief, etc.) More specifically, describe when you might or might not consider using the different priors. When do you think that they are “valid” or “invalid” in terms of representing your beliefs.

{Solution.}

The data-sets are very different and lead us to different prior assumptions. Doing this exercise was helpful to understand how different prior beliefs affects the design of our model and how it incorporates the data into the posterior distribution.

In order to better evaluate the results obtained, we will plot the 95% Credible Intervals for Mean and Standard Deviation for all priors and data-sets.

As we will see, these intervals presents differences due to our assumptions and the data-set used to calculate the posterior.

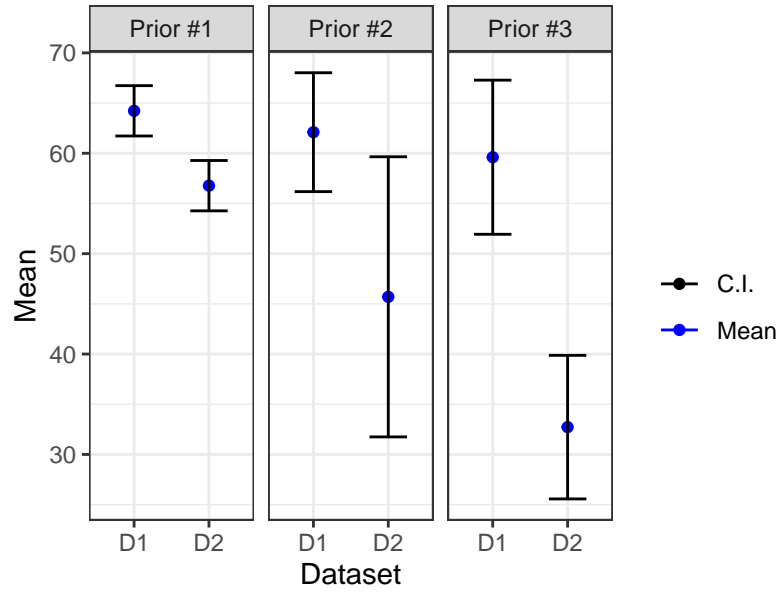


Figure 4: Comparison 95% Credible Intervals for Mean

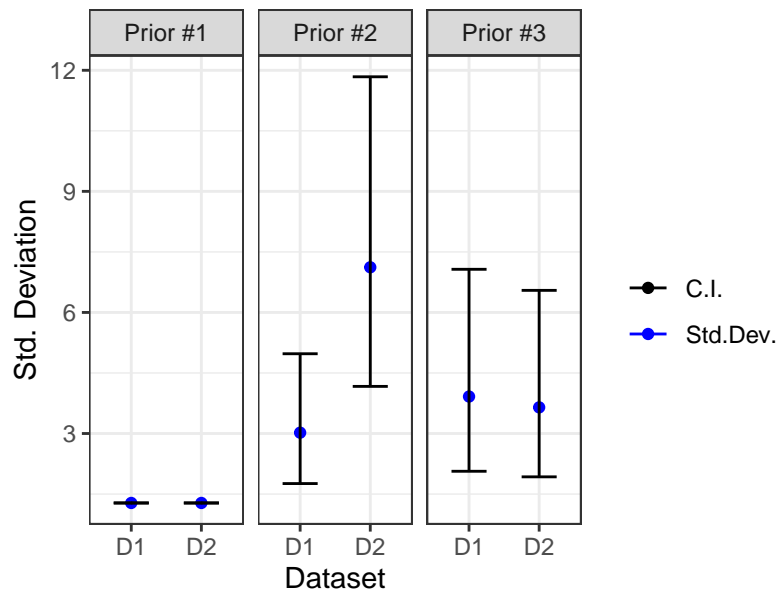


Figure 5: Comparison 95% Credible Intervals for Std.Deviation

For all priors we assumed a common belief that the average height is  $\mu_0 = 66$ .

In **Prior #1**, we also assumed known precision which was applied to both data-sets. We could see that the belief of  $\tau_0 = \frac{4}{9}$  says that we don't expect the data spreads out too much from the mean and represents the greater precision between all priors. For *D1* this might be fairly true but it doesn't apply when we see *D2*. The consequence is that the 95% Credible Interval using the posterior distribution are distributed around 66, which is inside *D1* data range but its far outside from the range of values of *D2*'s. Additionally we noticed the amplitude of those C.I.s are quite small reflecting the precision assumed by Prior #1.

In the case of **Prior #2** we have a more realistic situation where we don't know the precision  $\tau$  so we will have to estimate it through a *Normal-Gamma* model with additional parameters  $\theta$ ,  $\alpha$  and  $\beta$ . By setting our additional beliefs that the parameter  $\theta = 4$  (i.e., greater than 1) we assumed *more variability of our data* or - equivalently - less precision of  $\mu|\tau = t$ . We also assumed  $\alpha = 1$  and  $\beta = 36$  which made our posterior incorporate more weight from the data when estimating the precision of our model. That's why Prior #2 have larger credible intervals around the mean than Prior #1, but the fact of incorporating much of the data made our Credible Interval for Standard Deviation to be more elastic, specially for Data-set 2. This can also be seen on graph in item (c), where the posterior distribution (green line) is more flat in  $D2$  than in  $D1$ .

For **Prior #3** we reverted some of these effects by setting different values to our parameters:  $\theta = 0.1$ ,  $\alpha = 0.001$  and  $\beta = 0.001$ . The adjustment in  $\theta < 1$  and considered the values of  $X_i|\mu$  less spread out from the true mean which has made our model more *data-driven*, i.e., incorporated more information from our data. This is reflected in our credible intervals which has fitted better to our data.

In the same sense, smaller values for  $\alpha$  and  $\beta$  increased the influence of observed data into the posterior. What we see in this situation that both  $D1$  and  $D2$  were better adjusted to the posterior's credible interval for  $\mu$ . Additionally, the relationship  $\frac{\alpha}{\beta}$  equals to 1 (which is greater than in Prior #2) has made our belief of the precision is greater using this prior than the others. This can also be seen in the comparison of Credible Intervals for Standard Deviation which are approximately equal between both data-sets.

Our primary conclusion when comparing these 03 priors is that **Prior #3** seems to fit better to both observed sets of data.

## Question 3

(A basic, bare-bones, fixed effect meta analysis). Preamble: You are a scientist who is interested in the effect a class of drugs on the control of diabetes. One measure of diabetic control is called HbA1c. For diabetics who don't have good control, the value of HbA1c is too high. It is hoped that the drug lowers the value of HbA1c. For the purpose of this question, we will make a series of simplifying assumptions. They may not be realistic, but these assumptions make the problem manageable for a homework problem. At the end of this question, there are some comments which would be relevant to applying this question to a real world application.

Assumptions: (i) It will be assumed that there is one, universal value,  $\theta$ , which we are estimating. That is, the average amount HbA1c is lowered for all drugs in the class, for all dose levels, and for all populations. (ii) From a randomized trial  $i$ , one gets the average difference,  $\bar{Y}_i$ , and the standard error,  $S_i$ . The difference is between the value of the HbA1c at the end of the trial compared to the value at the beginning of the trial. The standard error is an estimate of the square root of the within trial variance of  $\bar{Y}_i$ . It will be assumed that the observed average difference,  $\bar{Y}_i$ , in trial  $i$  is conditionally normally distributed with mean  $\theta$  and variance  $S_i^2$  (given the value of  $\theta$  and  $S_i$ ). Note: here we are assuming that the sampled standard error is the true standard error. (So, measured without error). (iii) Also, before any experimental results are known, it is assumed that everyone's prior distribution for  $\theta$  is the vague, improper prior which is normal with mean zero and precision zero. (That is,  $\mu_0$  and  $\tau_0$  are both zero.)

Answer the following questions. Be sure to show the formula's that you used, don't just give numeric answers.

### item (a)

You do a randomized clinical trial with 209 subjects and you get a value for  $(\bar{Y}_1, S_1)$  of  $(-1.82, .21)$ . What is your (posterior) belief in  $\theta$  which is the average amount that the drug lowers the value of HbA1c? (This should be expressed as a distribution.)

{Solution.}

For this question, besides the assumptions stated in items (i), (ii) and (iii), we will also use the use the.

IN order to calculate the posterior distribution of  $\theta$  (the average amount that the drug lowers the value of HbA1c let's first identify the *Likelihood* and *Prior* distributions and apply them to Bayes Rule.

- The Likelihood distribution of  $\bar{Y}_1$  is

$$p(\bar{Y}_1|\theta, \tau) \sim N(\theta, \tau)$$

- ... with the Prior distribution<sup>3</sup> of  $\theta$  is

$$p(\theta|\mu_0, \tau_0) \sim N(0, 0), \text{ which is an improper prior.}$$

From the above definitions, it follows that the *posterior belief* of  $\theta$  can be expressed by the following posterior distribution:

$$p(\theta|\bar{Y}_1, S_1, \mu_0, \tau_0) \sim N(\theta'_0, \tau'_0) \quad (29)$$

... where  $\theta'_0$  and  $\tau'_0$  are given, respectively, by:

<sup>3</sup>Here we are considering the sampled Standard Error as the *true Standard Error*. By doing this, we don't need to model  $\tau$  as it is assumed known and it was measured with no error.



$$\theta'_0 = \frac{\mu_0 \tau_0 + \bar{Y}_1 S_1^{-2}}{\tau_0 + S_1^{-2}} \quad (30)$$

... and

$$\tau'_0 = \tau_0 + S_1^{-2} \quad (31)$$

Using the data from the experiment  $(\bar{Y}_1, S_1) = (-1.82, 0.21)$  and substituting it in (30) and (31) we obtained:

##

## Parameters of Posterior are: Theta0p= -1.820 and Tau0p= 22.6757

In this sense, applying these values in (29) we conclude that our belief of the amount that the drug lowers the value of  $HbA_{1c}$  is expressed by the following posterior distribution of  $\theta$ :

$$p(\theta | \bar{Y}_1, S_1, \mu_0, \tau_0) \sim N(-1.82, 22.6757) \quad (32)$$

### item (b)

After you ran your experiment, you learn that a colleague across the country has also run a clinical trial to measure the value of  $\theta$ . Your colleague had 79 subjects and got a value for  $(\bar{Y}_2, S_2)$  of  $(-1.02, .28)$ . Starting with the prior belief that you had from your own clinical trial, update your belief using the new data. (That is, use the information from the previous question as your prior distribution and then use this new information as data to get a new posterior distribution.) What is your new posterior belief on the values of  $\theta$ ?

{Solution.}

Now, considering that another study was carried out in other location (so, independently from the current study), we can consider this new data to calculate the new belief of  $\theta$  which is represented by the new posterior distribution and using our previous belief (i.e., the posterior obtained in item(a)) as our *new prior*.

In this sense, we now have the following:

- The Likelihood distribution of  $\bar{Y}_2$  is

$$p(\bar{Y}_2 | \theta, \tau) \sim N(\theta, \tau)$$

- ... with the Prior distribution of  $\theta$  is

$$p(\theta | \theta'_0, \tau'_0) \sim N(-1.82, 22.6757) \text{ - from (32)}$$

From the above definitions, it follows that the *updated belief of  $\theta$*  can be expressed by the following posterior distribution:

$$p(\theta | \bar{Y}_1, S_1, \bar{Y}_2, S_2, \theta'_0, \tau'_0) \sim N(\theta'_1, \tau'_1) \quad (33)$$

... where  $\theta'_1$  and  $\tau'_1$  are given, respectively, by:

$$\theta'_1 = \frac{\bar{Y}_1 S_1^{-2} + \bar{Y}_2 S_2^{-2}}{S_1^{-2} + S_2^{-2}} \quad (34)$$

... and

$$\tau_1' = S_1^{-2} + S_1^{-2} \quad (35)$$

Using the data from the new experiment  $(\bar{Y}_2, S_2) = (-1.02, 0.28)$  and substituting it in (34) and (35) we obtained:

##

## Parameters of Posterior are: Theta1p= -1.532 and Tau1p= 35.4308

In this sense, applying these values in (33) we conclude that our updated belief of the amount that the drug lowers the value of  $HbA_{1c}$  is expressed by the following posterior distribution of  $\theta$ :

$$p(\theta|\bar{Y}_1, S_1, \bar{Y}_2, S_2, \theta_0', \tau_0') \sim N(-1.532, 35.4308) \quad (36)$$

### item (c)

Now consider the problem from your colleagues point of view. When she first collects her data without seeing your data, what is her posterior belief for the value of  $\theta$ ? After she finds out about your data and she updates her belief, what is her new belief in the value of  $\theta$ ?

{Solution.}

Using the same approach from items (a) and (b) but with reversed order of data (i.e., considering the point of view of my colleague), we can calculate her the *first* and *updated* belief of  $\theta$  by doing:

1. Revert the order of data, i.e., set  $(\bar{Y}_1, S_1) = (-1.02, 0.28)$  and  $(\bar{Y}_2, S_2) = (-1.82, 0.21)$ ;
2. Apply formulas (29),(30) and (31) to calculate her *first belief* of  $\theta$ ;
3. Apply formulas (33),(34) and (35) to calculate her *updated belief* of  $\theta$

After following these steps we obtained the following results for each posterior:

##

## Parameters of Posterior are: Theta0p= -1.020 and Tau0p= 12.7551

##

## Parameters of Posterior are: Theta1p= -1.532 and Tau1p= 35.4308

**My colleague's First Belief of  $\theta$**

$$p(\theta|\bar{Y}_1, S_1, \mu_0, \tau_0) \sim N(-1.02, 12.7551) \quad (37)$$

**My colleague's Updated Belief of  $\theta$**

$$p(\theta|\bar{Y}_1, S_1, \bar{Y}_2, S_2, \theta_0', \tau_0') \sim N(-1.532, 35.4308) \quad (38)$$

We noticed that our updated beliefs on (36) and (38) converged to the same posterior distribution, given both pairs  $(\bar{Y}_1, S_1)$  and  $(\bar{Y}_2, S_2)$ .

**item (d)**

A few months later, you go to a conference and find several other labs who have study the problem and ran clinical trials. You learn that they got values for  $(n, \bar{Y}_i, S_i)$  of the following:  $(19, -1.90, 0.945)$ ,  $(100, -2.00, 0.285)$  and  $(20, -1.21, 0.545)$ . (Note:  $n$  is the number of subjects for each trial.) Pooling all this information, what is your new belief in the value of  $\theta$ ? Provide the general formula for combining this information. (Note: since you can easily get the formula by *goggling* up “fixed effect meta analysis”, there will be little weight to the actual formula. You will be mostly graded on explaining how the formula follows from the principles/formulas presented in the class material.)

Some closing notes. These notes are not part of the homework question. The data presented here is a subset of the data presented by Hirst et al, 2013, Diabetologia, 56 : 973 – 984. They considered a random effect meta analysis model. That model can be fit with some of the techniques that we will learn in a few weeks. For the purpose of this questions, the fixed effect model is simpler and highlights the connection between a simple Bayesian model and the main formula used in meta analysis. In reality, the random effect is probably more applicable in general. Otherwise, one is assuming that drug effect is constant over a wide range of conditions such as the different specific drugs in the drug family, the different doses used in a particular study, the different illness severity in the trial population as well as other trial population effects such as different gender mix, culture, social economic status, genetic mix. Also, as for assumptions used in this problem, the assumption that the trial standard error could be treated as the true standard error in the model is probably okay if there is a sufficiently large trial population. When doing a meta analysis, there are other considerations that one should consider such as publication bias. There have been a number of guidelines developed to perform a proper meta analysis study. If one is going to one for real, then one should consult the literature for these guidelines.

{Solution.}

The previous items showed we can derive a general formula to incorporate the data from other studies into a unique consolidated *belief of the average amount that the drug lowers the value of HbA<sub>1c</sub>*.

Considering we have  $K$  results from studies containing information of each pair  $(\bar{Y}_i, S_i)$  with  $i = 1, \dots, K$ , our posterior belief of  $\theta$  is distributed:

$$p(\theta | \bar{Y}_1, S_1, \dots, \bar{Y}_K, S_K, \theta'_0, \tau'_0, \dots, \theta'_{K-1}, \tau'_{K-1}) \sim N(\theta'_K, \tau'_K) \quad (39)$$

... where

$$\theta'_K = \frac{\sum_{i=1}^K \bar{Y}_i S_i^{-2}}{\sum_{i=1}^K S_i^{-2}} \quad (40)$$

... and

$$\tau'_K = \sum_{i=1}^K S_i^{-2} \quad (41)$$

Applying (39), (40) and (41) to the following data  $(19, -1.9, 0.945)$ ,  $(100, -2.0, 0.285)$ ,  $(20, -1.21, 0.545)$  we obtained the following results for the *overall belief* of  $\theta$ :

##

## Parameters of Posterior are: Theta5p= -1.62945 and Tau5p= 52.2288

$$p(\theta|\bar{Y}_1, S_1, \dots, \bar{Y}_5, S_5, \theta'_0, \tau'_0, \dots, \theta'_4, \tau'_4) \sim N(-1.62945, 52.2288) \quad (42)$$

In order to better understand how our beliefs change as we add data to calculate its posterior distribution, we plotted the 95% C.I. for  $\theta$  with the respective weighted standard error from each posterior. The x-axis represent how much data is contained on each posterior, for instance, “Y1” contains only  $\bar{Y}_1$ ; “Y1toY2” contains  $\bar{Y}_1$  and  $\bar{Y}_2$ ; “Y1toY3” contains  $\bar{Y}_1$ ,  $\bar{Y}_2$  and  $\bar{Y}_3$  and so on.

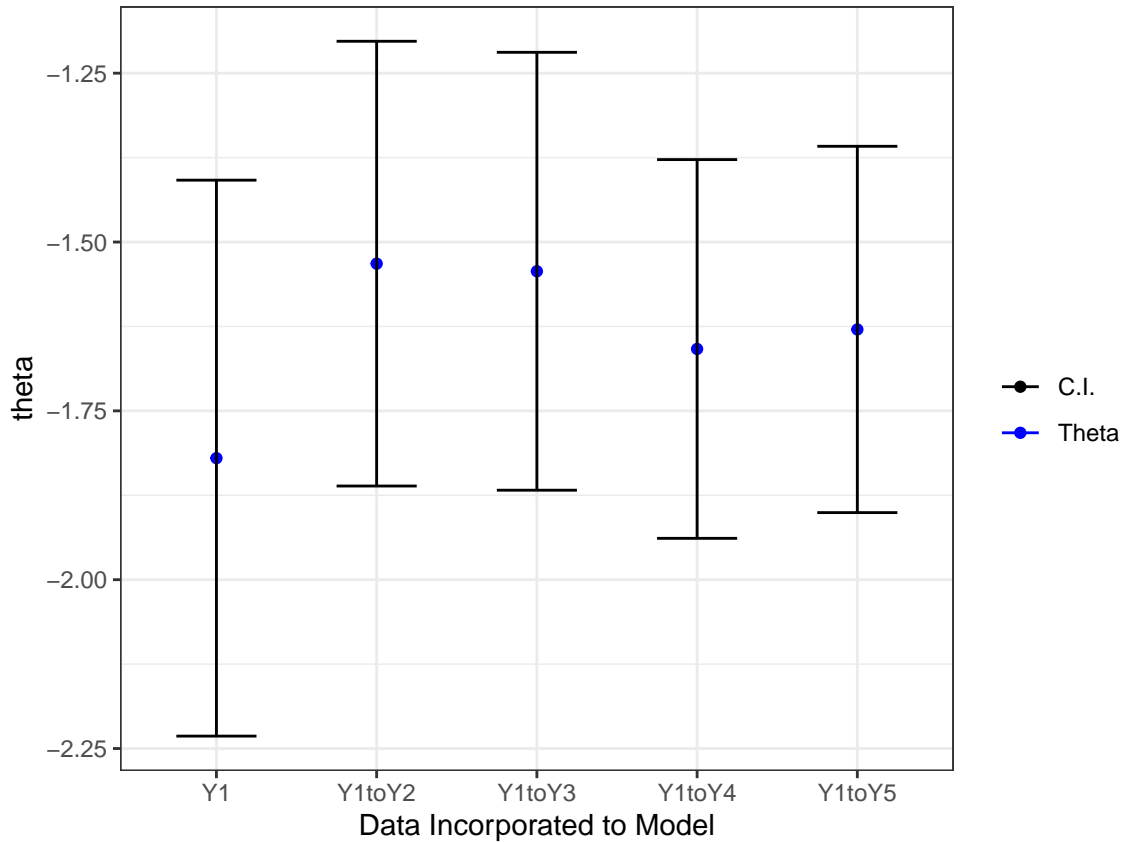


Figure 6: Comparison 95% Credible Intervals for our Belief as we add Data

We can see the adjustment of our belief (here represented by the centered blue point) and the corresponding credible interval. As we move from left-to-right, we increase the amount of data used to compose our posterior distribution of  $\theta$  and, as consequence, the more data we have, more adjusted will be our beliefs. The range of each interval also adjusts its size accordingly as the weighted standard error becomes smaller.

Additionally, we simulated each posterior to see this behavior represented by the density changes as we add data to our model.

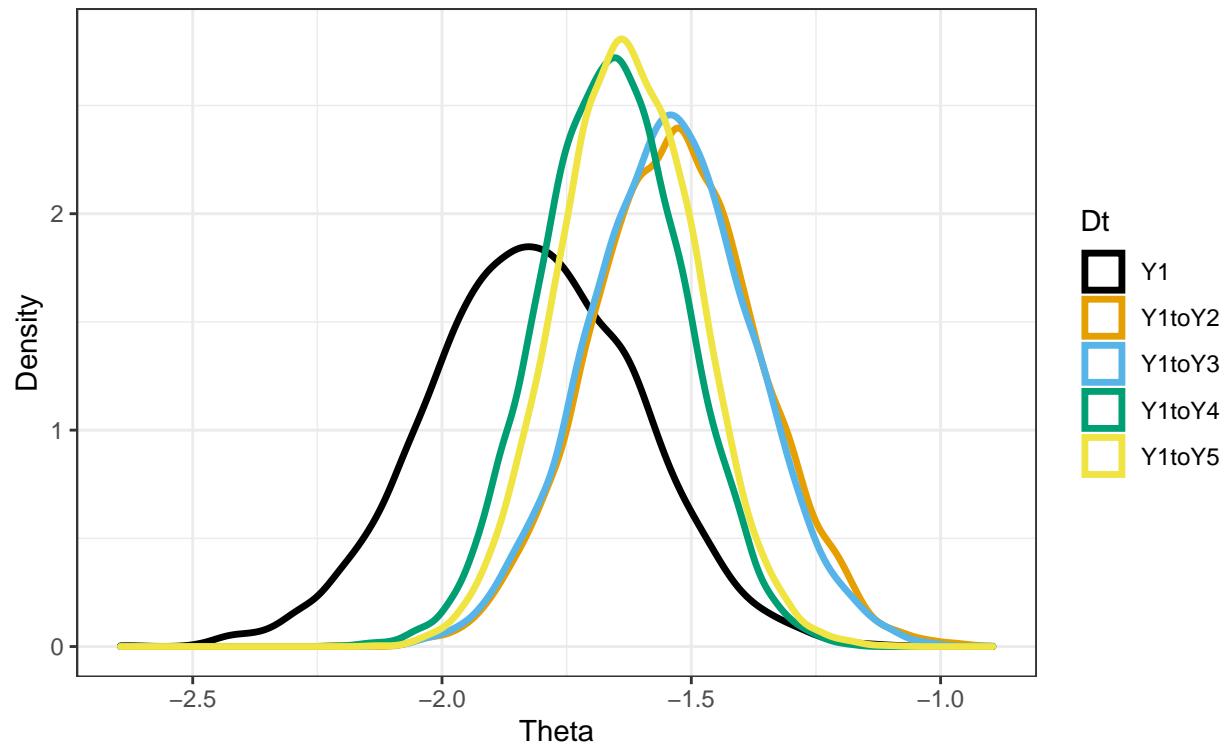


Figure 7: Sample Densities for each Posterior Belief of Theta

We can see the distribution moving from left (black curve) to the right (oranges and light-blue curves) and finally adjust to a more centered belief (green and yellow curves), both less flatten when compared with our initial belief with only  $\bar{Y}_1$ .

As conclusion, this exercise has shown how we can benefit from the the Bayesian approach to Fixed Effect Meta-Analysis by making use of the posterior framework and interpret the parameter of interest in terms of its probabilities.

Additionally, the process enabled the possibility to include prior beliefs, external sources of information, additional results obtained on other researches in order to refine our beliefs and predictions which is very useful in many situations.

## References

- [1] Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B. *Bayesian Data Analysis*. pp.34 - CRC Press, 2014.
- [2] Marin, J. M., Robert, C. *Bayesian Essentials with R*, 2nd. Ed. pp.49 - Springer, 2014.
- [3] DeGroot, M. H. *Optimal Statistical Decisions*, 1st. Ed. pp.164-171 - Wiley Interscience - 1970,2004.
- [4] Hirst J. A., Farmer A. J., Dyar A., Lung T. W. C., Stevens, R. J. *Estimating the effect of sulfonylurea on HbA<sub>1c</sub> in diabetes: a systematic review and meta-analysis*, Diabetologia 56:973–984 (2013)

- [5] Jansen, J.P., Crawford, B., Bergman, G., Stam, W. *Bayesian Meta-Analysis of Multiple Treatment Comparisons: An Introduction to Mixed Treatment Comparisons*, Value in Health, Vol.II, 5:956–964 (2008)
- [6] Smith, T.C., Spiegelhalter, D.J., Thomas, A. *Bayesian Approach to Random-Effects Meta-Analysis: A Comparative Study*, Statistics in Medicine, Vol. 14, pp.2685-2699, (1995)
- [7] Sutton A.J., Abrams K.R., Jones D.R., *Methods for Meta-Analysis in Medical Research*. London: Wiley, 2000.

## Appendix - R-Code

```
#---
#title: "CHL5223 - Applied Bayesian Methods - Assignment 1"
#author: "Luis Correia - Student No. 1006508566"
#date: "January 18th 2020"
#---

library(ggplot2)
library(tidyverse)

#-----
# CASE 1 - Non-Informative Prior
#-----

alpha1 <- beta1 <- 1.0
set.seed(125)
hist(rbeta(10000,1, 1),
     xlab = expression(list(theta[i]) ~ "|" ~ list(alpha[i],beta[i])))

#-----
# CASE 2 - Informative Prior
#-----
# --- Bisection Algorithm to obtain Alpha and Beta that satisfies the Informative Prior
#-----

epsilon <- 10(-6) # Precision desired for the estimates of alpha and beta

beg <- 0.000001 # Initial Value for Alpha
end <- 60.0 # Final value for Alpha
len <- 100 # Number of sections in the interval
Err <- FALSE

while(end - beg > epsilon) {
  aa<- seq(beg, end,length=len)
  bb<- aa
  tb <- cbind(aa, bb, cdfp95pct=pbeta(.60, aa,bb)-pbeta(.40,aa,bb))
  newbeg <- tail(which(tb[,3]<0.95),1)
  newend <- head(which(tb[,3]>0.95),1)
  if (newbeg == len) {
    # Bad choice of Beg and End - Stop condition
    cat("Input range to estimate Alpha is invalid.(Bisection Algorithm aborted!)" )
    beg <- end
    Err <- TRUE
  }
  else {
    beg <- tb[newbeg,1]
    end <- tb[newend,1]
  }
}

if (!Err) {
  alpha2 <- beta2 <- (beg+end)/2;
  names(alpha2) <- "Alpha"
  names(beta2) <- "Beta"
  cat("The prior distribution of Theta is Beta(", alpha2, ",", beta2,")\n")
  cat("This distribution guarantees that 95% of its mass is between 0.40 and 0.60 as we can see below:\n")
}
```

```

cat("> pbeta(0.4, 47.29982, 47.29982)=",pbeta(0.4, alpha2, beta2),"and ")
cat("pbeta(0.6, 47.29982, 47.29982)=",pbeta(0.6, alpha2, beta2))
cat("\n> qbeta(c(0.025,0.975), 47.29982, 47.29982)=",qbeta(c(0.025,0.975), 47.29982, 47.29982))

}
set.seed(1245)
hist(rbeta(10000,47.29982, 47.29982),
     xlab = expression(list(theta[i]) ~ "|" ~ list(alpha[i],beta[i])))

# Setup Variables
N1 <- N2 <- N3 <- 5001      # No. of voters per district
ThrsVictoryD1 <- trunc(N1/2,0)+1 # Threshold for victory on district #1
ThrsVictoryD2 <- trunc(N2/2,0)+1 # Threshold for victory on district #2
ThrsVictoryD3 <- trunc(N3/2,0)+1 # Threshold for victory on district #3

# Sample for Vote Intention on each district
SD1 <- data.frame(Purple = 53, Brown = 45);
SD1 <- cbind(District = 1, SD1, Total=sum(SD1))

SD2 <- data.frame(Purple = 72, Brown = 78);
SD2 <- cbind(District = 2, SD2, Total=sum(SD2))

SD3 <- data.frame(Purple = 18, Brown = 22);
SD3 <- cbind(District = 3, SD3, Total=sum(SD3))

kableExtra::kable(rbind(SD1, SD2, SD3),"latex", booktabs = T,
                  caption = "Sampling of Vote Intention by District")

#-----
# CASE 1 - Non-Informative Prior
#-----

# District 1 - Calculate Posterior Distribution, Mean, Std Deviation and 95% C.I.
posD1_Alpha_NI <- alpha1+SD1$Purple; #posD1_Alpha_NI
posD1_Beta_NI <- beta1+SD1$Total-SD1$Purple; #posD1_Beta_NI

cat("\n\n--- Summary for District 1 ---\n")

meanD1 <- posD1_Alpha_NI/(posD1_Alpha_NI+posD1_Beta_NI);
cat("\nPosterior Probability for the percent o voters for Purple party in Dist.#1 is",
    format(meanD1, digits = 6, nsmall = 3))

cat("\nMean=", format(meanD1, digits = 6, nsmall = 3))

stdevD1 <- sqrt((posD1_Alpha_NI*posD1_Beta_NI)/
               ((posD1_Alpha_NI+posD1_Beta_NI)^2*(posD1_Alpha_NI+posD1_Beta_NI+1)));
cat("\nStd. Deviation=", format(stdevD1, digits = 6, nsmall = 3))

ci_D1_NI <- qbeta(c(.025, .975),posD1_Alpha_NI,posD1_Beta_NI)
cat("\n95% Credible Region (", format(ci_D1_NI[1], digits = 4, nsmall = 3), ",",
    format(ci_D1_NI[2], digits = 4, nsmall = 3),")")

ProbD1_NI <- 1-extraDistr::pbbinom(ThrsVictoryD1-1,N1,posD1_Alpha_NI,

```



```

                                posD1_Beta_NI) # Beta-Binomial probability
cat("\nPosterior Probability Purple party win in Dist.#1 is",
    format(ProbD1_NI, digits = 6, nsmall = 3))

#-----
# CASE 1 - Non-Informative Prior
#-----

# District 2 - Calculate Posterior Distribution, Mean, Std Deviation and 95% C.I.
posD2_Alpha_NI <- alpha1+SD2$Purple; #posD2_Alpha_NI
posD2_Beta_NI <- beta1+SD2$Total-SD2$Purple; #posD2_Beta_NI

cat("\n\n--- Summary for District 2 ---\n")

meanD2 <- posD2_Alpha_NI/(posD2_Alpha_NI+posD2_Beta_NI);
cat("\nPosterior Probability for the percent o voters for Purple party in Dist.#2 is",
    format(meanD2, digits = 6, nsmall = 3))

cat("\nMean=", format(meanD2, digits = 6, nsmall = 3))

stdevD2 <- sqrt((posD2_Alpha_NI*posD2_Beta_NI)/
                ((posD2_Alpha_NI+posD2_Beta_NI)^2*(posD2_Alpha_NI+posD2_Beta_NI+1)));
cat("\nStd. Deviation=", format(stdevD2, digits = 6, nsmall = 3))

ci_D2_NI <- qbeta(c(.025, .975),posD2_Alpha_NI, posD2_Beta_NI)
cat("\n95% Credible Region (", format(ci_D2_NI[1], digits = 4, nsmall = 3), ",",
    format(ci_D2_NI[2], digits = 4, nsmall = 3),")")

ProbD2_NI <- 1-extraDistr::pbbinom(ThrsVictoryD2-1,N2,
                                posD2_Alpha_NI,posD2_Beta_NI) # Beta-Binomial probability
cat("\nPosterior Probability Purple party win in Dist.#2 is",
    format(ProbD2_NI, digits = 6, nsmall = 3))

#-----
# CASE 1 - Non-Informative Prior
#-----

# District 3 - Calculate Posterior Distribution, Mean, Std Deviation and 95% C.I.
posD3_Alpha_NI <- alpha1+SD3$Purple; #posD3_Alpha_NI
posD3_Beta_NI <- beta1+SD3$Total-SD3$Purple; #posD3_Beta_NI

cat("\n\n--- Summary for District 3 ---\n")

meanD3 <- posD3_Alpha_NI/(posD3_Alpha_NI+posD3_Beta_NI);
cat("\nPosterior Probability for the percent o voters for Purple party in Dist.#3 is",
    format(meanD3, digits = 6, nsmall = 3))

cat("\nMean=", format(meanD3, digits = 6, nsmall = 3))

stdevD3 <- sqrt((posD3_Alpha_NI*posD3_Beta_NI)/
                ((posD3_Alpha_NI+posD3_Beta_NI)^2*(posD3_Alpha_NI+posD3_Beta_NI+1)));
cat("\nStd. Deviation=", format(stdevD3, digits = 6, nsmall = 3))

```

```

ci_D3_NI <- qbeta(c(.025, .975), posD3_Alpha_NI, posD3_Beta_NI)
cat("\n95% Credible Region (", format(ci_D3_NI[1], digits = 4, nsmall = 3), ",",
    format(ci_D3_NI[2], digits = 4, nsmall = 3), ")")

ProbD3_NI <- 1-extraDistr::pbbinom(ThrsVictoryD3-1, N3,
                                posD3_Alpha_NI, posD3_Beta_NI) # Beta-Binomial probability
cat("\nPosterior Probability Purple party win in Dist.#3 is",
    format(ProbD3_NI, digits = 6, nsmall = 3))

#-----
# CASE 2 - Informative Prior
#-----

# District 1 - Calculate Posterior Distribution, Mean, Std Deviation and 95% C.I.
posD1_Alpha <- alpha2+SD1$Purple; #posD1_Alpha
posD1_Beta <- beta2+SD1$Total-SD1$Purple; #posD1_Beta

cat("\n\n--- Summary for District 1 ---\n")

meanD1 <- posD1_Alpha/(posD1_Alpha+posD1_Beta);
cat("\nPosterior Probability for the percent o voters for Purple party in Dist.#1 is",
    format(meanD1, digits = 6, nsmall = 3))

cat("\nMean=", format(meanD1, digits = 6, nsmall = 3))

stdevD1 <- sqrt((posD1_Alpha*posD1_Beta)/
                ((posD1_Alpha+posD1_Beta)^2*(posD1_Alpha+posD1_Beta+1)));
cat("\nStd. Deviation=", format(stdevD1, digits = 6, nsmall = 3))

ci_D1 <- qbeta(c(.025, .975), posD1_Alpha, posD1_Beta)
cat("\n95% Credible Region (", format(ci_D1[1], digits = 4, nsmall = 3), ",",
    format(ci_D1[2], digits = 4, nsmall = 3), ")")

ProbD1 <- 1-extraDistr::pbbinom(ThrsVictoryD1-1, N1,
                                posD1_Alpha, posD1_Beta) # Beta-Binomial probability
cat("\nPosterior Probability Purple party win in Dist.#1 is",
    format(ProbD1, digits = 6, nsmall = 3))

#-----
# CASE 2 - Informative Prior
#-----

# District 2 - Calculate Posterior Distribution, Mean, Std Deviation and 95% C.I.
posD2_Alpha <- alpha2+SD2$Purple; #posD2_Alpha
posD2_Beta <- beta2+SD2$Total-SD2$Purple; #posD2_Beta

cat("\n\n--- Summary for District 2 ---\n")

meanD2 <- posD2_Alpha/(posD2_Alpha+posD2_Beta);
cat("\nPosterior Probability for the percent o voters for Purple party in Dist.#2 is",
    format(meanD2, digits = 6, nsmall = 3))

cat("\nMean=", format(meanD2, digits = 6, nsmall = 3))

```

```

stdevD2 <- sqrt((posD2_Alpha*posD2_Beta)/
               ((posD2_Alpha+posD2_Beta)^2*(posD2_Alpha+posD2_Beta+1)));
cat("\nStd. Deviation=", format(stdevD2, digits = 6, nsmall = 3))

ci_D2 <- qbeta(c(.025, .975),posD2_Alpha, posD2_Beta)
cat("\n95% Credible Region (", format(ci_D2[1], digits = 4, nsmall = 3), ",",
    format(ci_D2[2], digits = 4, nsmall = 3),")")

ProbD2 <- 1-extraDistr::pbbinom(ThrsVictoryD2-1,N2,
                              posD2_Alpha,posD2_Beta) # Beta-Binomial probability
cat("\nPosterior Probability Purple party win in Dist.#2 is",
    format(ProbD2, digits = 6, nsmall = 3))

#-----
# CASE 2 - Informative Prior
#-----

# District 3 - Calculate Posterior Distribution, Mean, Std Deviation and 95% C.I.
posD3_Alpha <- alpha2+SD3$Purple; #posD3_Alpha
posD3_Beta <- beta2+SD3$Total-SD3$Purple; #posD3_Beta

cat("\n\n--- Summary for District 3 ---\n")

meanD3 <- posD3_Alpha/(posD3_Alpha+posD3_Beta);
cat("\nPosterior Probability for the percent o voters for Purple party in Dist.#3 is",
    format(meanD3, digits = 6, nsmall = 3))

cat("\nMean=", format(meanD3, digits = 6, nsmall = 3))

stdevD3 <- sqrt((posD3_Alpha*posD3_Beta)/
               ((posD3_Alpha+posD3_Beta)^2*(posD3_Alpha+posD3_Beta+1)));
cat("\nStd. Deviation=", format(stdevD3, digits = 6, nsmall = 3))

ci_D3 <- qbeta(c(.025, .975),posD3_Alpha, posD3_Beta)
cat("\n95% Credible Region (", format(ci_D3[1], digits = 4, nsmall = 3), ",",
    format(ci_D3[2], digits = 4, nsmall = 3),")")

ProbD3 <- 1-extraDistr::pbbinom(ThrsVictoryD3-1,N3,
                              posD3_Alpha,posD3_Beta) # Beta-Binomial probability
cat("\nPosterior Probability Purple party win in Dist.#3 is",
    format(ProbD3, digits = 6, nsmall = 3))

# Set seed for DEBUG
set.seed(1965)

N=10000 # No. of Elections to Simulate

# Variable to store X_i (No. of votes for Purple Party on each district)
Sim_X1_NI <- rep(NA,N)
Sim_X2_NI <- rep(NA,N)
Sim_X3_NI <- rep(NA,N)

# Case 1 - Non-Informative Prior

```

```

for (n in 1:N) {
  # Generate a Theta_i for each district
  thetaD1 <- rbeta(1,posD1_Alpha_NI,posD1_Beta_NI);
  thetaD2 <- rbeta(1,posD2_Alpha_NI,posD2_Beta_NI)
  thetaD3 <- rbeta(1,posD3_Alpha_NI,posD3_Beta_NI)

  #Calculate The total No. of Votes for Purple Party on each district
  Sim_X1_NI[n] <- rbinom(1,N1,thetaD1)
  Sim_X2_NI[n] <- rbinom(1,N2,thetaD2)
  Sim_X3_NI[n] <- rbinom(1,N3,thetaD3)
}

# Plot the probabilities for each District - Case 1 - Non-Informative Prior

ProbPurpWinD1_NI <- length(which(Sim_X1_NI>(ThrsVictoryD1)))/N;
ProbPurpWinD2_NI <- length(which(Sim_X2_NI>(ThrsVictoryD2)))/N;
ProbPurpWinD3_NI <- length(which(Sim_X3_NI>(ThrsVictoryD3)))/N;

# Calculating probabilities of win 2 districts

ProbPurpWinD1D2_NI <- length(which(Sim_X1_NI>(ThrsVictoryD1)
                                & Sim_X2_NI>(ThrsVictoryD2)))/N;
ProbPurpWinD1D3_NI <- length(which(Sim_X1_NI>(ThrsVictoryD1)
                                & Sim_X3_NI>(ThrsVictoryD3)))/N;
ProbPurpWinD2D3_NI <- length(which(Sim_X2_NI>(ThrsVictoryD2)
                                & Sim_X3_NI>(ThrsVictoryD3)))/N;

# Calculating probabilities of win 3 districts
ProbPurpWinD1D2D3_NI <- length(which(Sim_X1_NI>(ThrsVictoryD1) &
                                    Sim_X2_NI>(ThrsVictoryD2) &
                                    Sim_X3_NI>(ThrsVictoryD3)))/N;

SD11 <- cbind(District = 1, Freq=length(which(Sim_X1_NI>(ThrsVictoryD1))),
              Prop=ProbPurpWinD1_NI)
SD12 <- cbind(District = 2, Freq=length(which(Sim_X2_NI>(ThrsVictoryD2))),
              Prop=ProbPurpWinD2_NI)
SD13 <- cbind(District = 3, Freq=length(which(Sim_X3_NI>(ThrsVictoryD3))),
              Prop=ProbPurpWinD3_NI)

kableExtra::kable(rbind(SD11, SD12, SD13),"latex", booktabs = T,
                  caption = "Purple Party wins in 10,000 simulations")

SD21 <- cbind(District = "1 & 2",
              Freq=length(which(Sim_X1_NI>(ThrsVictoryD1) &
                              Sim_X2_NI>(ThrsVictoryD2))),
              Prop=ProbPurpWinD1D2_NI)
SD22 <- cbind(District = "1 & 3",
              Freq=length(which(Sim_X1_NI>(ThrsVictoryD1) &
                              Sim_X3_NI>(ThrsVictoryD3))),
              Prop=ProbPurpWinD1D3_NI)
SD23 <- cbind(District = "2 & 3",
              Freq=length(which(Sim_X2_NI>(ThrsVictoryD2) &
                              Sim_X3_NI>(ThrsVictoryD3))),

```

```

Prop=ProbPurpWinD2D3_NI)

kableExtra::kable(rbind(SD21, SD22, SD23),"latex", booktabs = T,
  caption = "Purple Party wins at least 02 districts in 10,000 simulations")

#cat("\nProbability of Purple Win District #1 and #2 =",ProbPurpWinD1D2_NI)
#cat("\nProbability of Purple Win District #1 and #3 =",ProbPurpWinD1D3_NI)
#cat("\nProbability of Purple Win District #2 and #3 =",ProbPurpWinD2D3_NI)

F1 <- as.integer(SD21[1,2])-length(which(Sim_X1_NI>(ThrsVictoryD1) &
  Sim_X2_NI>(ThrsVictoryD2) &
  Sim_X3_NI>(ThrsVictoryD3)))
F2 <- as.integer(SD22[1,2])-length(which(Sim_X1_NI>(ThrsVictoryD1) &
  Sim_X2_NI>(ThrsVictoryD2) &
  Sim_X3_NI>(ThrsVictoryD3)))
F3 <- as.integer(SD23[1,2])-length(which(Sim_X1_NI>(ThrsVictoryD1) &
  Sim_X2_NI>(ThrsVictoryD2) &
  Sim_X3_NI>(ThrsVictoryD3)))
F4 <- length(which(Sim_X1_NI>(ThrsVictoryD1) & Sim_X2_NI>(ThrsVictoryD2) &
  Sim_X3_NI>(ThrsVictoryD3)))

SD31 <- cbind(District = "Only 1 & 2", Freq=F1, Prop=F1/N)
SD32 <- cbind(District = "Only 1 & 3", Freq=F2, Prop=F2/N)
SD33 <- cbind(District = "Only 2 & 3", Freq=F3, Prop=F3/N)
SD34 <- cbind(District = "All", Freq=F4, Prop=F4/N)

kableExtra::kable(rbind(SD31, SD32, SD33, SD34),"latex",
  booktabs = T,
  caption = "Purple Party wins exactly 02 districts / All districts in 10,000 simulations")

# Plot the probabilities for each District - Case 1 - Non-Informative Prior

cat("\nProbability of Purple Win District #1=",ProbPurpWinD1_NI)
cat("\nProbability of Purple Win District #2=",ProbPurpWinD2_NI)
cat("\nProbability of Purple Win District #3=",ProbPurpWinD3_NI)

# Determining intersections with Full Victory

cat("\n\nProbability that the Purple Party have majority in Town Council=",sum(F1+F2+F3+F4)/N)

# Variable to store X_i (No. of votes for Purple Party on each district)
Sim_X1 <- rep(NA,N)
Sim_X2 <- rep(NA,N)
Sim_X3 <- rep(NA,N)

# Case 1 - Non-Informative Prior
for (n in 1:N) {
  # Generate a Theta_i for each district
  thetaD1 <- rbeta(1,posD1_Alpha,posD1_Beta)
  thetaD2 <- rbeta(1,posD2_Alpha,posD2_Beta)
  thetaD3 <- rbeta(1,posD3_Alpha,posD3_Beta)

  #Calculate The total No. of Votes for Purple Party on each district

```

```

Sim_X1[n] <- rbinom(1,N1,thetaD1)
Sim_X2[n] <- rbinom(1,N2,thetaD2)
Sim_X3[n] <- rbinom(1,N3,thetaD3)
}

# Plot the probabilities for each District - Case 1 - Non-Informative Prior

ProbPurpWinD1 <- length(which(Sim_X1>(ThrsVictoryD1)))/N;
ProbPurpWinD2 <- length(which(Sim_X2>(ThrsVictoryD2)))/N;
ProbPurpWinD3 <- length(which(Sim_X3>(ThrsVictoryD3)))/N;

# Calculating probabilities of win 2 districts

ProbPurpWinD1D2 <- length(which(Sim_X1>(ThrsVictoryD1) &
                               Sim_X2>(ThrsVictoryD2)))/N;
ProbPurpWinD1D3 <- length(which(Sim_X1>(ThrsVictoryD1) &
                               Sim_X3>(ThrsVictoryD3)))/N;
ProbPurpWinD2D3 <- length(which(Sim_X2>(ThrsVictoryD2) &
                               Sim_X3>(ThrsVictoryD3)))/N;

# Calculating probabilities of win 3 districts
ProbPurpWinD1D2D3 <- length(which(Sim_X1>(ThrsVictoryD1) &
                                  Sim_X2>(ThrsVictoryD2) &
                                  Sim_X3>(ThrsVictoryD3)))/N;

SD11 <- cbind(District = 1,
              Freq=length(which(Sim_X1>(ThrsVictoryD1))),
              Prop=ProbPurpWinD1)
SD12 <- cbind(District = 2,
              Freq=length(which(Sim_X2>(ThrsVictoryD2))),
              Prop=ProbPurpWinD2)
SD13 <- cbind(District = 3,
              Freq=length(which(Sim_X3>(ThrsVictoryD3))),
              Prop=ProbPurpWinD3)

kableExtra::kable(rbind(SD11, SD12, SD13),"latex", booktabs = T,
                  caption = "Purple Party wins in 10,000 simulations")

SD21 <- cbind(District = "1 & 2",
              Freq=length(which(Sim_X1>(ThrsVictoryD1) &
                               Sim_X2>(ThrsVictoryD2))),
              Prop=ProbPurpWinD1D2)
SD22 <- cbind(District = "1 & 3",
              Freq=length(which(Sim_X1>(ThrsVictoryD1) &
                               Sim_X3>(ThrsVictoryD3))),
              Prop=ProbPurpWinD1D3)
SD23 <- cbind(District = "2 & 3",
              Freq=length(which(Sim_X2>(ThrsVictoryD2) &
                               Sim_X3>(ThrsVictoryD3))),
              Prop=ProbPurpWinD2D3)

kableExtra::kable(rbind(SD21, SD22, SD23),"latex", booktabs = T,
                  caption = "Purple Party wins at least 02 districts in 10,000 simulations")

```

```

F1 <- as.integer(SD21[1,2])-length(which(Sim_X1>(ThrsVictoryD1)
                                         & Sim_X2>(ThrsVictoryD2) &
                                         Sim_X3>(ThrsVictoryD3)))
F2 <- as.integer(SD22[1,2])-length(which(Sim_X1>(ThrsVictoryD1)
                                         & Sim_X2>(ThrsVictoryD2) &
                                         Sim_X3>(ThrsVictoryD3)))
F3 <- as.integer(SD23[1,2])-length(which(Sim_X1>(ThrsVictoryD1)
                                         & Sim_X2>(ThrsVictoryD2) &
                                         Sim_X3>(ThrsVictoryD3)))
F4 <- length(which(Sim_X1>(ThrsVictoryD1) & Sim_X2>(ThrsVictoryD2) &
                  Sim_X3>(ThrsVictoryD3)))

SD31 <- cbind(District = "Only 1 & 2", Freq=F1, Prop=F1/N)
SD32 <- cbind(District = "Only 1 & 3", Freq=F2, Prop=F2/N)
SD33 <- cbind(District = "Only 2 & 3", Freq=F3, Prop=F3/N)
SD34 <- cbind(District = "All", Freq=F4, Prop=F4/N)

kableExtra::kable(rbind(SD31, SD32, SD33, SD34),"latex", booktabs = T,
                    caption = "Purple Party wins exactly 02 districts / All districts in 10,000 simulation")

# Plot the probabilities for each District - Case 1 - Non-Informative Prior

cat("\nProbability of Purple Win District #1=",ProbPurpWinD1)
cat("\nProbability of Purple Win District #2=",ProbPurpWinD2)
cat("\nProbability of Purple Win District #3=",ProbPurpWinD3)

# Determining intersections with Full Victory

cat("\n\nProbability that the Purple Party have majority in Town Council=",sum(F1+F2+F3+F4)/N)

kableExtra::kable(rbind(cbind(District="1",Simul=ProbPurpWinD1_NI,
                              BetaBin=format(ProbD1_NI, digits = 4, nsmall = 4)),
                        cbind(District="2",Simul=ProbPurpWinD2_NI,
                              BetaBin=format(ProbD2_NI, digits = 4, nsmall = 4)),
                        cbind(District="3",Simul=ProbPurpWinD3_NI,
                              BetaBin=format(ProbD3_NI, digits = 4, nsmall = 4))),
                  "latex", booktabs = T,
                  caption = "Simulated vs. and Beta-Binomial Probability (Case 1: Non-Informative Prior)")

kableExtra::kable(rbind(cbind(District="1",Simul=ProbPurpWinD1,
                              BetaBin=format(ProbD1, digits = 4, nsmall = 4)),
                        cbind(District="2",Simul=ProbPurpWinD2,
                              BetaBin=format(ProbD2, digits = 4, nsmall = 4)),
                        cbind(District="3",Simul=ProbPurpWinD3,
                              BetaBin=format(ProbD3, digits = 4, nsmall = 4))),
                  "latex", booktabs = T,
                  caption = "Simulated vs. and Beta-Binomial Probability (Case 2: Informative Prior)")

# Plot Sample densities to Compare Non-Informative vs. Informative Prior
D_Work <- rbind(data.frame(Distr1=Sim_X1_NI,Distr2=Sim_X2_NI,Distr3=Sim_X3_NI,PType="Non-Informative"),
               data.frame(Distr1=Sim_X1,Distr2=Sim_X2,Distr3=Sim_X3,PType="Informative"))
#colnames(D_Work) <- c("Distr1","Distr2","Distr3","PType" )
D_Work %>%

```



```

pivot_longer(cols=Distr1:Distr3, names_to = "District", values_to = "Votes" ) %>%
ggplot(mapping = aes(x = Votes, group = District))+
geom_density(aes(colour=District), size=1.2)+ #linetype=District,
labs(x = "Votes", y = "Density") +
facet_grid(~PType)+
theme_bw()

TbCI <- data.frame(
  District=c(1, 2, 3),
  NonInf=c(format(ci_D1_NI[2]-ci_D1_NI[1], digits = 4, nsmall = 4),
            format(ci_D2_NI[2]-ci_D2_NI[1], digits = 4, nsmall = 4),
            format(ci_D3_NI[2]-ci_D3_NI[1], digits = 4, nsmall = 4)),
  Inform=c(c(format(ci_D1[2]-ci_D1[1], digits = 4, nsmall = 4),
              format(ci_D2[2]-ci_D2[1], digits = 4, nsmall = 4),
              format(ci_D3[2]-ci_D3[1], digits = 4, nsmall = 4))))

kableExtra::kable(TbCI, "latex", booktabs = T,
                  caption = "Comparison of Amplitude of 95\\% Credible Interval")

# The Data
#-----
D1 <- c(53,49,63,72,55,65)
D2 <- c(28,27,36,42,25,35)

# Prior 1 - Calculate the Posterior
#-----
# Belief 0 - Tau is known
tau <- 1/36

# Belief 1 - Average height is 66 inches
mu0 <- 66

# Belief 3 - Mu0 is between 63 and 69 with 95% probability
tau0 <- 4/9

# Calculate the Posterior for each data-set
#-----
# Data-set 1 -
cat("-----")
cat("\nSummary Statistics for Data-Set1- (",D1,")\n")
mu_l1 <- (tau0*mu0+length(D1)*tau*mean(D1))/(tau0+length(D1)*tau);
cat("\nPosterior Mean is ",format(mu_l1, digits = 6, nsmall = 3))

tau_l1 <- (tau0+length(D1)*tau);
cat("\nPosterior Std. Deviation is ",format(sqrt(1/tau_l1), digits = 6, nsmall = 3))

ci_DS1 <- qnorm(c(.025, .975),mean=mu_l1, sd=sqrt(1/tau_l1))
cat("\n95% Credible Region for mean (", format(ci_DS1[1], digits = 4, nsmall = 3), ",",
      format(ci_DS1[2], digits = 4, nsmall = 3),")")

#xP1D1 <- rnorm(N,mean=mu_l1,sd=sqrt(1/tau_l1))
#cat("\n\nBy simulation (N=",N,") >>> Mean=",mean(xP1D1)," / Std.Dev.=",sd(xP1D1))

```



```

# Data-set 2 -
cat("\n\n-----")
cat("\nSummary Statistics for Data-Set2- (" ,D2," )\n")
mu_l2 <- (tau0*mu0+length(D2)*tau*mean(D2))/(tau0+length(D2)*tau);
cat("\nPosterior Mean is ",format(mu_l2, digits = 6, nsmall = 3))

tau_l2 <- (tau0+length(D2)*tau);
cat("\nPosterior Std. Deviation is ",format(sqrt(1/tau_l2), digits = 6, nsmall = 3))

ci_DS2 <- qnorm(c(.025, .975),mean=mu_l2, sd=sqrt(1/tau_l2))
cat("\n95% Credible Region for mean (", format(ci_DS2[1], digits = 4, nsmall = 3), ",",
    format(ci_DS2[2], digits = 4, nsmall = 3),")")

#xP1D2 <- rnorm(N,mean=mu_l2,sd=sqrt(1/tau_l2))
#cat("\n\nBy simulation (N=",N," ) >>> Mean=",mean(xP1D2)," | Std.Dev.=",sd(xP1D2))

# Prior 2 - Calculate the Posterior
#-----
# Belief 1 - Average height is 66 inches
mu0 <- 66

# Belief 2 - Tau is Gamma(alpha,beta)
thetaP2 <- 4
alphaP2 <- 1
betaP2 <- 36

# Calculate the Posterior for each data-set
#-----
# Data-set 1 -
cat("-----")
cat("\nSummary Statistics for Data-Set1- (" ,D1," )\n")
alpha_l1P2 <- alphaP2+length(D1)/2
beta_l1P2 <- betaP2+0.5*sum((D1-mean(D1))^2)+
    (thetaP2*length(D1)*(mean(D1)-mu0)^2)/(2*(thetaP2+length(D1)))

## Analytically calculating the statistics

mu_l1P2 <- (thetaP2*mu0+length(D1)*mean(D1))/(thetaP2+length(D1))
cat("\nPosterior Mean is ",format(mu_l1P2, digits = 6, nsmall = 3))

tau_l1P2A <- 1/((beta_l1P2/((thetaP2+length(D1))*alpha_l1P2))*
    ((2*alpha_l1P2)/(2*alpha_l1P2-2)))
cat("\n(Analytically)Posterior Std. Deviation is ",
    format(sqrt(1/tau_l1P2A), digits = 6, nsmall = 3)) # Analytically

TauG <- rgamma(10*N,alpha_l1P2, beta_l1P2)
m1 <- rnorm(10*N, mean=mu_l1P2, sd=1/sqrt(TauG*(thetaP2+length(D1))))
tau_l1P2S <- 1/var(m1)
cat("\n(Simulated)Posterior Std. Deviation is ",
    format(sqrt(1/tau_l1P2S), digits = 6, nsmall = 3))

tau_l1P2 <- tau_l1P2S # Using the Simulated output to be compatible w/ C.I.

```

```

ci_DS1P2 <- qnorm(c(.025, .975),mean=mu_11P2, sd=sqrt(1/tau_11P2))
cat("\n95% Credible Region for mean (",
    format(ci_DS1P2[1], digits = 4, nsmall = 3),
    ",",format(ci_DS1P2[2], digits = 4, nsmall = 3),")")

ci_DS1P2SD <- quantile(1/sqrt(TauG*(thetaP2+length(D1))), c(.025, .975))
cat("\n95% Credible Region for Std.Deviation (",
    format(ci_DS1P2SD[1], digits = 4, nsmall = 3),
    ",",format(ci_DS1P2SD[2], digits = 4, nsmall = 3),")")

#-----
# Data-set 2 -
cat("\n\n-----")
cat("\nSummary Statistics for Data-Set2- (" ,D2,")\n")
alpha_12P2 <- alphaP2+length(D2)/2
beta_12P2 <- betaP2+0.5*sum((D2-mean(D2))^2)+
    (thetaP2*length(D2)*(mean(D2)-mu0)^2)/(2*(thetaP2+length(D2)))

## Analytically calculating the statistics

mu_12P2 <- (thetaP2*mu0+length(D2)*mean(D2))/(thetaP2+length(D2))
cat("\nPosterior Mean is ",format(mu_12P2, digits = 6, nsmall = 3))

#--- NEW
tau_12P2A <- 1/((beta_12P2/((thetaP2+length(D2))*alpha_12P2))*
    ((2*alpha_12P2)/(2*alpha_12P2-2)))
cat("\n(Analytically) Posterior Std. Deviation is ",
    format(sqrt(1/tau_12P2A), digits = 6, nsmall = 3)) # Analytically

TauG <- rgamma(10*N,alpha_12P2, beta_12P2)
m2 <- rnorm(10*N, mean=mu_12P2, sd=1/sqrt(TauG*(thetaP2+length(D2))))
tau_12P2S <- 1/var(m2)
cat("\n(Simulated) Posterior Std. Deviation is ",
    format(sqrt(1/tau_12P2S), digits = 6, nsmall = 3))

tau_12P2 <- tau_12P2S # Using the Simulated output to be compatible w/ C.I.
#----

ci_DS2P2 <- qnorm(c(.025, .975),mean=mu_12P2, sd=sqrt(1/tau_12P2))
cat("\n95% Credible Region for mean (",
    format(ci_DS2P2[1], digits = 4, nsmall = 3),
    ",",format(ci_DS2P2[2], digits = 4, nsmall = 3),")")

#--- NEW
ci_DS2P2SD <- quantile(1/sqrt(TauG*(thetaP2+length(D2))), c(.025, .975))
cat("\n95% Credible Region for Std.Deviation (",
    format(ci_DS2P2SD[1], digits = 4, nsmall = 3),
    ",",format(ci_DS2P2SD[2], digits = 4, nsmall = 3),")")
#---

# Prior 3 - Calculate the Posterior
#-----
# Belief 1 - Average height is 66 inches

```

```

mu0 <- 66

# Belief 2 - Tau is Gamma(alpha,beta)
thetaP3 <- 0.1
alphaP3 <- 0.001
betaP3 <- 0.001

# Calculate the Posterior for each data-set
#-----
# Data-set 1 -
cat("-----")
cat("\nSummary Statistics for Data-Set1- (",D1,")\n")
alpha_l1P3 <- alphaP3+length(D1)/2
beta_l1P3 <- betaP3+0.5*sum((D1-mean(D1))^2)+
  (thetaP3*length(D1)*(mean(D1)-mu0)^2)/(2*(thetaP3+length(D1)))

## Analytically calculating the statistics

mu_l1P3 <- (thetaP3*mu0+length(D1)*mean(D1))/(thetaP3+length(D1))
cat("\nPosterior Mean is ",format(mu_l1P3, digits = 6, nsmall = 3))

tau_l1P3A <- 1/((beta_l1P3/((thetaP3+length(D1))*alpha_l1P3))*
  ((2*alpha_l1P3)/(2*alpha_l1P3-2))) # v18
# tau_l1P3A <- 1/(alpha_l1P3/beta_l1P3^2) # v19
cat("\n(Analytically) Posterior Std. Deviation is ",
  format(sqrt(1/tau_l1P3A), digits = 6, nsmall = 3)) # Analytically

TauG <- rgamma(10*N,alpha_l1P3, beta_l1P3);
m1 <- rnorm(10*N, mean=mu_l1P3, sd=1/sqrt(TauG*(thetaP3+length(D1))))
tau_l1P3S <- 1/var(m1)
cat("\n(Simulated) Posterior Std. Deviation is ",
  format(sqrt(1/tau_l1P3S), digits = 6, nsmall = 3)) # v18

tau_l1P3 <- tau_l1P3S # Using the Simulated output to be compatible w/ C.I.

ci_DS1P3 <- qnorm(c(.025, .975),mean=mu_l1P3, sd=sqrt(1/tau_l1P3))
cat("\n95% Credible Region for mean (", format(ci_DS1P3[1], digits = 4, nsmall = 3),
  ",",format(ci_DS1P3[2], digits = 4, nsmall = 3),")")

ci_DS1P3SD <- quantile(1/sqrt(TauG*(thetaP3+length(D1))), c(.025, .975))
cat("\n95% Credible Region for Std.Deviation (",
  format(ci_DS1P3SD[1], digits = 4, nsmall = 3),
  ",",format(ci_DS1P3SD[2], digits = 4, nsmall = 3),")")

#-----
# Data-set 2 -
cat("\n\n-----")
cat("\nSummary Statistics for Data-Set2- (",D2,")\n")
alpha_l2P3 <- alphaP3+length(D2)/2
beta_l2P3 <- betaP3+0.5*sum((D2-mean(D2))^2)+
  (thetaP3*length(D2)*(mean(D2)-mu0)^2)/(2*(thetaP3+length(D2)))

## Analytically calculating the statistics

```

```

mu_l2P3 <- (thetaP3*mu0+length(D2)*mean(D2))/(thetaP3+length(D2))
cat("\nPosterior Mean is ",format(mu_l2P3, digits = 6, nsmall = 3))

#--- NEW
tau_l2P3A <- 1/((beta_l2P3/((thetaP3+length(D2))*alpha_l2P3))*
               ((2*alpha_l2P3)/(2*alpha_l2P3-2)))
cat("\n(Analytically)Posterior Std. Deviation is ",
    format(sqrt(1/tau_l2P3A), digits = 6, nsmall = 3)) # Analytically

TauG <- rgamma(10*N,alpha_l2P3, beta_l2P3)
m2 <- rnorm(10*N, mean=mu_l2P3, sd=1/sqrt(TauG*(thetaP3+length(D2))))
tau_l2P3S <- 1/var(m2)
cat("\n(Simulated)Posterior Std. Deviation is ",
    format(sqrt(1/tau_l2P3S), digits = 6, nsmall = 3))

tau_l2P3 <- tau_l2P3S # Using the Simulated output to be compatible w/ C.I.
#----

ci_DS2P3 <- qnorm(c(.025, .975),mean=mu_l2P3, sd=sqrt(1/tau_l2P3))
cat("\n95% Credible Region for mean (",
    format(ci_DS2P3[1], digits = 4, nsmall = 3),
    ",",format(ci_DS2P3[2], digits = 4, nsmall = 3),")")

#--- NEW
ci_DS2P3SD <- quantile(1/sqrt(TauG*(thetaP3+length(D2))), c(.025, .975))
cat("\n95% Credible Region for Std.Deviation (",
    format(ci_DS2P3SD[1], digits = 4, nsmall = 3),
    ",",format(ci_DS2P3SD[2], digits = 4, nsmall = 3),")")
#---

# Plot Sample densities to Compare Non-Informative vs. Informative Prior

xP1D1 <- rnorm(N,mean=mu_l1,sd=sqrt(1/tau_l1))
xP1D2 <- rnorm(N,mean=mu_l2,sd=sqrt(1/tau_l2))
xP2D1 <- rnorm(N,mean=mu_l1P2,sd=sqrt(1/((thetaP2+length(D1))*
                                     rgamma(N,alpha_l1P2,beta_l1P2))))
xP2D2 <- rnorm(N,mean=mu_l2P2,sd=sqrt(1/((thetaP2+length(D2))*
                                     rgamma(N,alpha_l2P2,beta_l2P2))))
xP3D1 <- rnorm(N,mean=mu_l1P3,sd=sqrt(1/((thetaP3+length(D1))*
                                     rgamma(N,alpha_l1P3,beta_l1P3))))
xP3D2 <- rnorm(N,mean=mu_l2P3,sd=sqrt(1/((thetaP3+length(D2))*
                                     rgamma(N,alpha_l2P3,beta_l2P3))))

D_Work <- as.tibble(rbind(data.frame(Height=xP1D1,Prior="1",Dataset="D1"),
                          data.frame(Height=xP1D2,Prior="1",Dataset="D2"),
                          data.frame(Height=xP2D1,Prior="2",Dataset="D1"),
                          data.frame(Height=xP2D2,Prior="2",Dataset="D2"),
                          data.frame(Height=xP3D1,Prior="3",Dataset="D1"),
                          data.frame(Height=xP3D2,Prior="3",Dataset="D2")))

#D_Work %>%
# ggplot(aes(x = Height))+
# geom_density(aes(x=Height,group=Prior, colour=Prior), size=1.2)+
# labs(x = "Height", y = "Density") +

```

```

# scale_linetype_manual(name = "Dataset", values = c("solid", "dashed"))+
# facet_grid(~Dataset)+
# theme_bw()

D_Work %>%
  ggplot(aes(x = Height))+
  geom_density(aes(x=Height, linetype=Dataset, colour=Prior), size=1.2)+
  labs(x = "Height", y = "Density") +
  scale_linetype_manual(values =c("solid", "dashed"))+
  theme_bw()

# We will use the MArginal distribution of mu (t-Student as described)

cat("-----")
cat("\nSummary Statistics of Predictive Distribution\n")

cat("\n#>>> Analytically ---\n")
cat("\nPosterior Mean is ",format(mu_11P3, digits = 6, nsmall = 3)) # Revise V19

VarNewX <- (beta_11P3/((thetaP3+length(D1))*alpha_11P3))*
  ((2*alpha_11P3)/(2*alpha_11P3-2)) # From Supplemental paper

cat("\nPosterior Std. Deviation is ",format(sqrt(VarNewX), digits = 6, nsmall = 3))

# Analytically (v19)
ci_NewX <- qt(c(.025, .975),
  2*alpha_11P3)*sqrt(beta_11P3/((thetaP3+length(D1))*alpha_11P3))+mu_11P3
cat("\n95% Credible Region for New Observation (",
  format(ci_NewX[1], digits = 4, nsmall = 3),
  ",",format(ci_NewX[2], digits = 4, nsmall = 3),")")

# Simulating from t Distribution # Revised v19
cat("\n\n#>>> Simulated ---\n")
NewX <- rt(N,2*alpha_11P3)*sqrt(beta_11P3/((thetaP3+length(D1))*alpha_11P3))+mu_11P3

cat("\nPosterior Mean is ",format(mean(NewX), digits = 6, nsmall = 3)) # V18
cat("\nPosterior Std. Deviation is ",format(sd(NewX), digits = 6, nsmall = 3))
# cat("\n\n",quantile(NewX, c(.025, .975)), "\n\n")
ci_NewX <- quantile(NewX, c(.025, .975))
cat("\n95% Credible Region for New Observation (",
  format(ci_NewX[1], digits = 4, nsmall = 3),
  ",",format(ci_NewX[2], digits = 4, nsmall = 3),")")

ggplot(data=data.frame(Height=NewX),aes(x = Height))+
  geom_density(aes(x=Height), size=1.2, color="deepskyblue")+
  labs(x = "Height", y = "Density") +
  theme_bw()

# Build Dataframes to display C.I.s

TbCID1 <- data.frame(Prior = c("Prior #1","Prior #2","Prior #3"),
  Min = c(ci_DS1[1],ci_DS1P2[1],ci_DS1P3[1]),
  Max= c(ci_DS1[2],ci_DS1P2[2],ci_DS1P3[2]),

```

```

MM = c(mu_l1,mu_l1P2,mu_l1P3),
DS = c("D1","D1","D1"))

TbCID2 <- data.frame(Prior = c("Prior #1","Prior #2","Prior #3"),
  Min = c(ci_DS2[1],ci_DS2P2[1],ci_DS2P3[1]),
  Max= c(ci_DS2[2],ci_DS2P2[2],ci_DS2P3[2]),
  MM = c(mu_l2,mu_l2P2,mu_l2P3),
  DS = c("D2","D2","D2"))

TbCID1SD <- data.frame(Prior = c("Prior #1","Prior #2","Prior #3"),
  Min = c(1/sqrt(tau_l1), ci_DS1P2SD[1],ci_DS1P3SD[1]),
  Max= c(1/sqrt(tau_l1), ci_DS1P2SD[2],ci_DS1P3SD[2]),
  MM = c(1/sqrt(tau_l1), 1/sqrt(tau_l1P2),1/sqrt(tau_l1P3)),
  DS = c("D1", "D1", "D1"))

TbCID2SD <- data.frame(Prior = c("Prior #1","Prior #2","Prior #3"),
  Min = c(1/sqrt(tau_l2), ci_DS2P2SD[1],ci_DS2P3SD[1]),
  Max= c(1/sqrt(tau_l2), ci_DS2P2SD[2],ci_DS2P3SD[2]),
  MM = c(1/sqrt(tau_l2), 1/sqrt(tau_l2P2),1/sqrt(tau_l2P3)),
  DS = c("D2", "D2", "D2"))

# Plot C.I.s
ggplot(data = rbind(TbCID1,TbCID2)) +
  geom_point(aes(x=DS, y=MM, color = "Mean")) +
  geom_errorbar(aes(x=DS, ymin=Min, ymax=Max,
    color = "C.I."), width = 0.5) +
  xlab("Dataset") +
  ylab("Mean") +
  scale_color_manual(name = "",
    values = c("Mean" = "Blue",
      "C.I." = "Black"))+
  facet_grid(~Prior)+
  theme_bw()

# Plot C.I.s
ggplot(data = rbind(TbCID1SD,TbCID2SD)) +
  geom_point(aes(x=DS, y=MM, color = "Std.Dev.")) +
  geom_errorbar(aes(x=DS, ymin=Min, ymax=Max,
    color = "C.I."), width = 0.5) +
  xlab("Dataset") +
  ylab("Std. Deviation") +
  scale_color_manual(name = "",
    values = c("Std.Dev." = "Blue",
      "C.I." = "Black"))+
  facet_grid(~Prior)+
  theme_bw()

# Calculus of ThetaOp and TauOp

# Sampled Data
Ybar1 <- -1.82
S1 <- .21
n1 <- 209

```

```

thetaOp <- (Ybar1*S1^(-2))/(S1^(-2))
tauOp <- (S1^(-2))

cat("\nParameters of Posterior are: ThetaOp=",
    format(thetaOp, digits = 6, nsmall = 3),
    " and TauOp=",format(tauOp, digits = 6, nsmall = 3))

# Calculus of ThetaIp and TauIp

# New Sampled Data
Ybar2 <- -1.02
S2 <- .28
n2 <- 79

thetaIp <- (Ybar1*S1^(-2)+Ybar2*S2^(-2))/(S1^(-2)+S2^(-2))
tauIp <- (S1^(-2)+S2^(-2))

cat("\nParameters of Posterior are: ThetaIp=",
    format(thetaIp, digits = 6, nsmall = 3),
    " and TauIp=",format(tauIp, digits = 6, nsmall = 3))

# Calculus of ThetaOp and TauOp

# Sampled Data in reversed order
thetaOp <- (Ybar2*S2^(-2))/(S2^(-2))
tauOp <- (S2^(-2))

cat("\nParameters of Posterior are: ThetaOp=",
    format(thetaOp, digits = 6, nsmall = 3),
    " and TauOp=",format(tauOp, digits = 6, nsmall = 3))

# Calculus of ThetaIp and TauIp
# New Sampled Data from p.o.v. of my colleague

thetaIp <- (Ybar2*S2^(-2)+Ybar1*S1^(-2))/(S2^(-2)+S1^(-2))
tauIp <- (S2^(-2)+S1^(-2))

cat("\nParameters of Posterior are: ThetaIp=",
    format(thetaIp, digits = 6, nsmall = 3),
    " and TauIp=",format(tauIp, digits = 6, nsmall = 3))

# General Formula to calculate posterior belief of Theta_K

# Setup variables
DataHbA1c <- data.frame (
  Ybar = c(Ybar1, Ybar2, -1.9, -2.0, -1.21),
  S = c(S1, S2, 0.945, 0.285, 0.545)
)
K = nrow(DataHbA1c)

# Calculate Overall Theta_Kp

theta_p <- rep(0,K)

```

```

tau_p <- rep(0,K)

pCrit <- qnorm(.975,0,1)

# Calculating the Posterior parameters for each cumulative entry
theta_p <- cumsum((DataHbA1c$Ybar[1:K]*DataHbA1c$S[1:K]^(-2))/
  cumsum(DataHbA1c$S[1:K]^(-2)))
tau_p <- cumsum(DataHbA1c$S[1:K]^(-2))

# Calculating the 95% C.I. for each belief of theta
ci_p = data.frame (
  Dt = c("Y1","Y1toY2","Y1toY3","Y1toY4","Y1toY5"),
  Min = theta_p-sqrt(1/tau_p)*pCrit,
  MM = theta_p,
  Max = theta_p+sqrt(1/tau_p)*pCrit
)

cat("\nParameters of Posterior are: Theta5p=",
  format(theta_p[K], digits = 6, nsmall = 3),
  " and Tau5p=",format(tau_p[K], digits = 6, nsmall = 3))

# Plot C.I.s
ggplot(data = ci_p) +
  geom_point(aes(x=Dt, y=MM, color = "Theta")) +
  geom_errorbar(aes(x=Dt, ymin=Min, ymax=Max,
    color = "C.I."), width = 0.5) +
  xlab("Data Incorporated to Model") +
  ylab(expression("theta")) +
  scale_color_manual(name = "",
    values = c("Theta" = "Blue",
      "C.I." = "Black"))+
  theme_bw()

# Simulating The change in the density of posterior belief as we add data
Y1 <- rnorm(N,mean=theta_p[1],sd=sqrt(1/tau_p[1]))
Y1toY2 <- rnorm(N,mean=theta_p[2],sd=sqrt(1/tau_p[2]))
Y1toY3 <- rnorm(N,mean=theta_p[3],sd=sqrt(1/tau_p[3]))
Y1toY4 <- rnorm(N,mean=theta_p[4],sd=sqrt(1/tau_p[4]))
Y1toY5 <- rnorm(N,mean=theta_p[5],sd=sqrt(1/tau_p[5]))

# Plot Sample densities to Compare Non-Informative vs. Informative Prior
D_WorkHbA1c <- as.tibble(rbind(data.frame(Theta=Y1, Dt="Y1"),
  data.frame(Theta=Y1toY2,Dt="Y1toY2"),
  data.frame(Theta=Y1toY3,Dt="Y1toY3"),
  data.frame(Theta=Y1toY4,Dt="Y1toY4"),
  data.frame(Theta=Y1toY5,Dt="Y1toY5")))

# The palette with black:
cbp1 <- c("#999999", "#E69F00", "#56B4E9", "#009E73",
  "#F0E442", "#0072B2", "#D55E00", "#CC79A7")
cbp2 <- c("#000000", "#E69F00", "#56B4E9", "#009E73",
  "#F0E442", "#0072B2", "#D55E00", "#CC79A7")
cbp3 <- c("#FFDB6D", "#C4961A", "#F4EDCA", "#D16103",
  "#C3D7A4", "#52854C", "#4E84C4", "#293352")

```



```
D_WorkHbA1c %>%  
  ggplot(aes(x = Theta))+  
  geom_density(aes(x=Theta,group=Dt, colour=Dt), size=1.2)+  
  labs(x = "Theta", y = "Density") +  
  scale_color_manual(values = cbp2)+  
  theme_bw()
```