Student Name : HILL JIGISHKUMAR MODI

Student ID : S3827516

I certify that this is all my own original work. If I took any parts from elsewhere, then they were non-essential parts of the assignment, and they are clearly attributed in my submission. I will show I agree to this honour code by typing "Yes": *Yes*.

ASSIGNMENT 1 – APPLIED BAYESIAN STATISTICS

Bayesian Analysis to find the Estimated Mean and Variance of the Sale Price of Properties of Melbourne with Informative and Non-Informative Priors and Check with Null Hypotheses Test

Hill Jigishkumar Modi

s3827516@student.rmit.edu.au

RMIT University, City Campus

Melbourne, Australia

Date: 30/08/2020

CONTENTS

INTRODUCTION

AIM / OBJECTIVE

The main goal of this assignment is to Estimate Mean sale price of properties of Melbourne and Variance of the Sale Price, by doing Bayesian Analysis.

We will use Gibbs Sampling method with implementation of MCMC (Monte Carlo Markov Chains). So that, we will understand how MCMC works with different parameters and values and we will check if the posterior distribution derived is acceptable or nor by using Diagnostic checking of MCMC such as MCMC Representativeness, MCMC Accuracy, MCMC Efficiency.

This report will also give us insights of how Bayesian Analysis is different than Classical Analysis and how high degree of belief (Informative Prior) or low degree of belief (Non – Informative Prior) impacts on posterior distribution.

MCMC METHODS

We have posterior distribution, but we don't know the expected value, variance, shape of posterior distribution and other characteristics.

MCMC methods generates data, they proposed values from that unknown distributions by using density.

In MCMC, current step is depended only on previous steps and not rest of the steps. So, we compose the chain, which consists of accepted values.

We want to generate data from target distributions, but we can't do that, because we don't know the actual mathematical formula of target distributions. But, we know the mathematical formula of proposed distribution.

So, we pick random value X from proposed distribution, put it into target distribution and get the propose density.

$$\mathbf{P_{move} = Density_p / Density_c}$$

If we repeat this process many times, in a long run, we will get the data points on Target or Posterior Distribution.

DESCRIPTIVE STATISTICS

Here, we have data about sale price of properties in Melbourne in CSV file called Assignment1PropertyPrices.csv.

We have column **"Sale Price (100K)"** which tells Sell Price.

For example, Sale Price 10 = 100,000 AUD.

Let, **X** be the Random Variable.

X : Sale Price of Properties in Melbourne in 100K

Here, we have Continuous Data.

SUMMARY

Summary

Filter

	Mean	Median	SD	Variance	Q1	Q3	IQR	Min	Max	Row
1	6.09	4.5	5.12	26.25	3.5	6.55	3.04962	2	70	10000

Figure 1

By doing, Descriptive summary in R, now we have better understanding of the data, so we can interpret the data better and tell overall story.

I find following.

- **Mean** – Mean sale price of Properties in Melbourne is 6.09K AUD.
- **Median** - Median sale price of Properties in Melbourne is 4.5K AUD,

And other important attributes such as **Quantile 1 = 3.5**, **Quantile 2 = 6.55**, **Minimum = 2**, **Maximum =70** and more.

VISUALIZATION

Visualization of data will always give us better understanding than written one.
Histogram of the same data, will give us some better understanding about the data.

HISTOGRAM

Figure 2

All the values of sale price have been plotted.

Most of the Sale price is between (0 K AUD – 20K AUD), while Sale price of (0 K AUD – 5 K AUD) have been highest.

Mean, Median, Max, Min have been also plotted.

Here, Maximum is on the value of 70 - indicates highest Sale Price is 70K AUD.

It depicts that maximum value is too far away from bars (Last bar in the range of (30K AUD – 35K AUD)). Which might rouse the case of outliers.

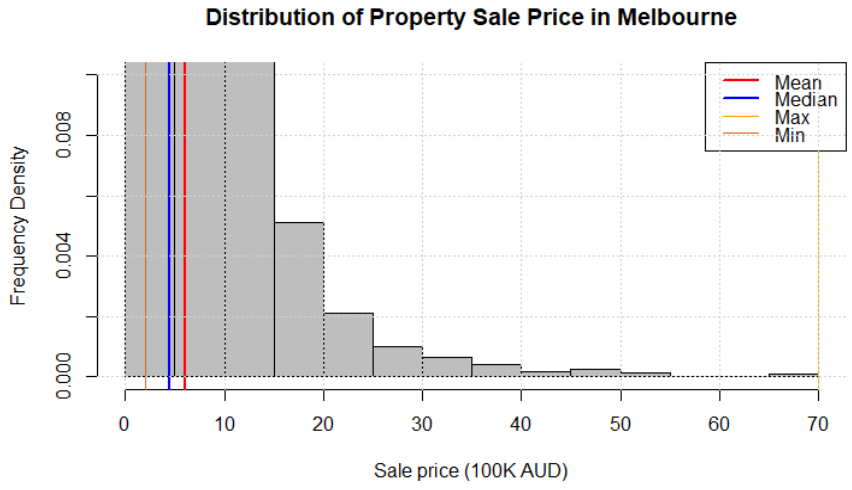

Figure 3

But, if we look closely, we will notice that, there are some properties, which have also been sold between (35K AUD – 70K AUD).

Even though, in this perspective, it can be depict that range of 55K – 65K have no values, if we again shorten the y-axe to the limit of (0,0.002). We will notice the presence of bars.

We will check outliers with the help of boxplot.

BOXPLOT

Figure 4

Here, from the boxplot, we can say that many values are outside of upper bound (11K) of boxplot – All these values are called as outliers in Statistics world, but as these are real data of Sales price of Properties – Removing them is not a good idea.

So, we will stuck with the data we have, without modifying anything.

Now, the next step is Mathematical Model

MATHEMATICAL MODEL

Here, we have instruction that Sale Price is distributed as Normal distribution.

X is our Random Variable, which indicates Sale Price, then

$X \sim Normal(\mu,\sigma^2)$,

Where, μ → Mean value of Sale Price(100K AUD) of Properties in the Melbourne.

σ^2 → Variance or uncertainty around the mean Sale Price(100K AUD) of properties of Melbourne.

Here, we will use **Gibbs Sampling Method**.

- Gibbs Sampling method is type of *Random Walk through the parameter space*.
- We have *two parameters* or *two success probabilities*, each of them is willing to meet each other in Random Walk.

As we have given instruction that,

- μ is distributed as Normal Distribution.

$\mu \sim Normal(\mu_0,\sigma_0^2)$,

Where, μ_0 → Prior information of Mean Sale Price (100K AUD) of properties of Melbourne,

σ_0^2 → Degree of belief in Prior Information of Mean Sale price.

Domain of Normal Distribution is (- ∞ < x < ∞).

- σ^2 is distributed as Gamma Distribution.

σ^2 ~ Gamma(α,β),

Where, α → Shape parameter,

β → Scale parameter.

To find α and β,

We have two equations from Gamma Distribution.

$\alpha = \frac{\mu^2}{\sigma^2}$ ----- (1), $\beta = \frac{\mu}{\sigma^2}$ -----(2)

Where, μ → prior information of variance or uncertainty in the Sale Price (100K AUD) in the properties of Melbourne,

σ^2 → degree of belief in prior information of Variance.

Domain of Gamma Distribution is (0 < x < ∞).

- So, We will use **Gibbs Sampling for Normal-Gamma Model.**

X ~ Normal(μ,σ^2),

μ ~ Normal(μ_0,σ_0^2),

σ^2 ~ Gamma(α,β),

SOME IMPORTANT TERMINOLOGY

MCMC REPRESENTATIVENESS

- **Burn-in Steps** = these are preliminary steps during which the chain moves from its unrepresentative initial values to the model region of the Posterior.
- **Shrink Factor =** It is also known as Gelman-Rubin Statistics or Potential Scale reduction Factor.
  - = If value is less than 1.2 → no problem.
  - = If value is 1.0 → they are fully converged.
  - = If value is greater than 1.2 →chains have not converged.

Trace plots, Density plots, Shrink Factor suggest whether Burn-in period has been suitably passed or not, chains are well mixed or not and representative of the posterior.

MCMC ACCURACY

Larger the sample, more stable and accurate Posterior Estimates and Related High Density Interval. (HDI).

- **Auto-correlation –** In Gibbs Sampling, we are taking successive steps to generate the chains, current position is somehow depended on previous and next position is somehow depended on current position. So, this correlation between successive steps is called **Autocorrelation**.

And, high or significant correlation is obviously not a good scenario.

- **Thinning –** Every k^{th} step in the chain will stored, and it will be breaking down the dependency between successive steps or points.

  If k = 10, then 1^{st},11^{th},21^{st},31^{st} … steps will be stored.

Thinning creates less Auto-correlated chains. Means, to overcome the problem of high auto-correlation, we can increase the number or thinning steps.

- **ESS (Effective Sample Size) =** Measure of Independent information left in the auto-correlated chains.

$$\textbf{ESS} = \frac{N}{1 + 2\sum_{k=1}^{\infty} ACF(K)},$$

Where, **ACF(K)** → auto-correlation of chains at Lag K.

For best scenario, ESS is closed to N, indicating there is nearly 0 auto-correlation. So, the goal is to push ESS as high as Possible.

- **Lag K =** Auto-correlation between current point and K previous point.
- **MCSE (Monte Carlo Standard Error)** = $\frac{\text{SD}}{\text{ESS}}$;

Where, **SD** → Standard Deviations.

ESS and MCSE suggests how stable and accurate our chains are.

NON-INFORMATIVE PRIOR

Non – Informative prior means - very low degree of belief, huge variance.

TRIAL AND ERROR

SPECIFY THE PRIOR INFORMATION

- Here, we have given that *mean sale price is 750,000 AUD (7.5 K AUD)*, Mean is distributed as Normal Distribution.

  As this Mean sale price is non-informative, we have very low degree of belief.

  When we have very low degree of belief, we have huge variance and histogram plot is also less concentrated.

  So, I've put the value of $\mu_0 = 7.5$ as Mean of normal prior distribution of population mean

  $\sigma_0^2 = 1$ as Variance of normal prior distribution of population mean

- We have also given that *Standard Deviation is 600,000 AUD (6K AUD)*

  Which means, variance is 36K, with very low degree of belief.

  So, $\mu = 36$

  $\sigma^2 = 1$,

But, variance is distributed as Gamma Distribution, so we have to find α and β from formula (1) and (2).

$$\alpha = \frac{\mu^2}{\sigma^2} = \frac{36^2}{1} = 1296,$$

$$\beta = \frac{\mu}{\sigma^2} = \frac{36}{1} = 36,$$

And put them in the app.

Next, I've put following values to the different Parameters.

- **Comparison value for the Population mean** = 6
- **Comparison Value for the population variance** = 27
- **Number of chain** = 3
- **Number of Burn-in Steps** = 50
- **Number of Saved Steps** = 500
- **Number of thinning steps** = 2

OUTPUT-ANALYSIS

The output is, estimated mean = 6.11 and Estimated Variance = 27.7

POSTERIOR DISTRIBUTION

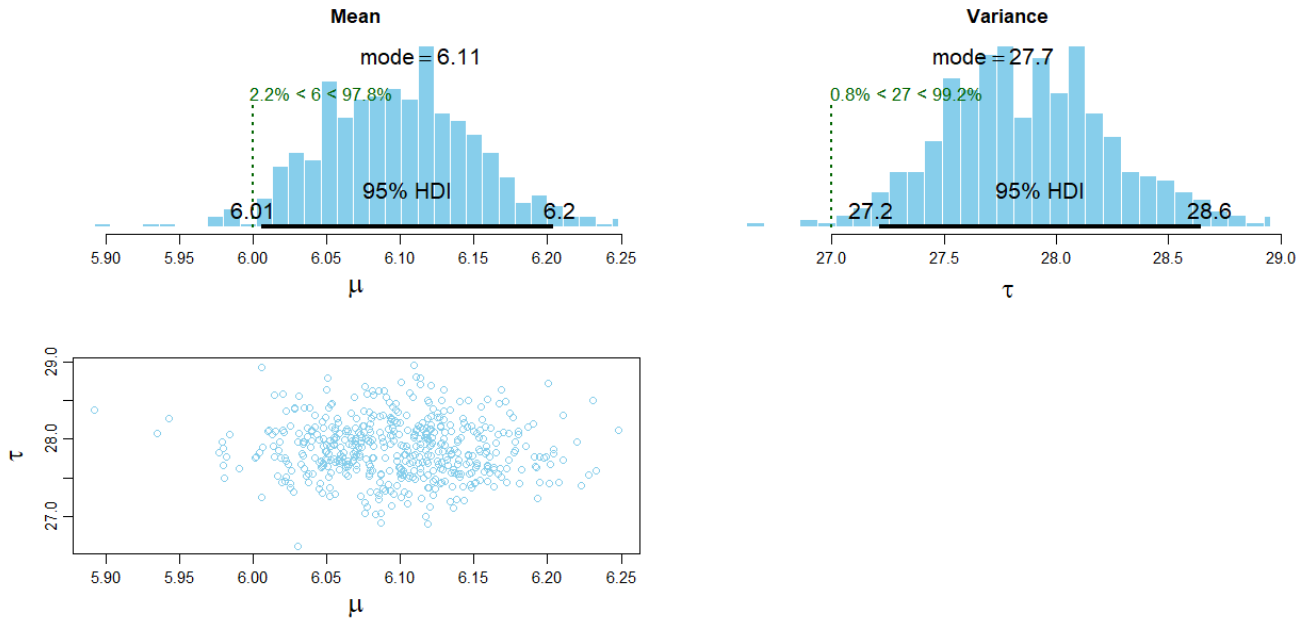

Figure 5

But, when I check the diagnostics of both mean and the variance, I found something unusual. Furthermore, histograms of both are also not symmetric.

Here, histograms have gaps, therefore don't give us exact idea about density of success probability.

Mode is the Bayesian estimate of success probability.

DIAGNOSTIC OF MEAN, μ

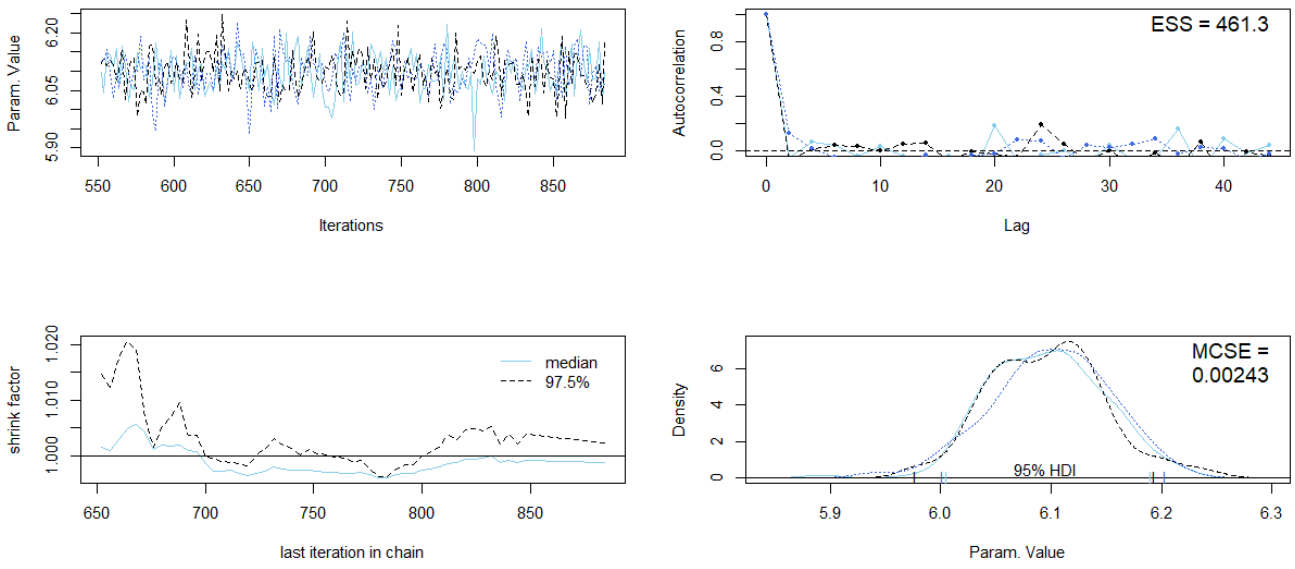

Figure 6

Here, trace plot is fluctuating from (5.9 to 6.2), which is obviously not the good case, so here, MCMC Representativeness diagnostic is not satisfied even though Shrink Factor in within acceptable range. So, we can overcome this problem by rising the number of Burn-in Steps and Saved Steps.

In Density plots, chains are not properly overlapped, which is directly related to the ESS that means, there is high auto-correlation, which is not significant, we can overcome this problem by increasing the number of thinning steps.

DIAGNOSTIC OF VARIANCE, σ^2

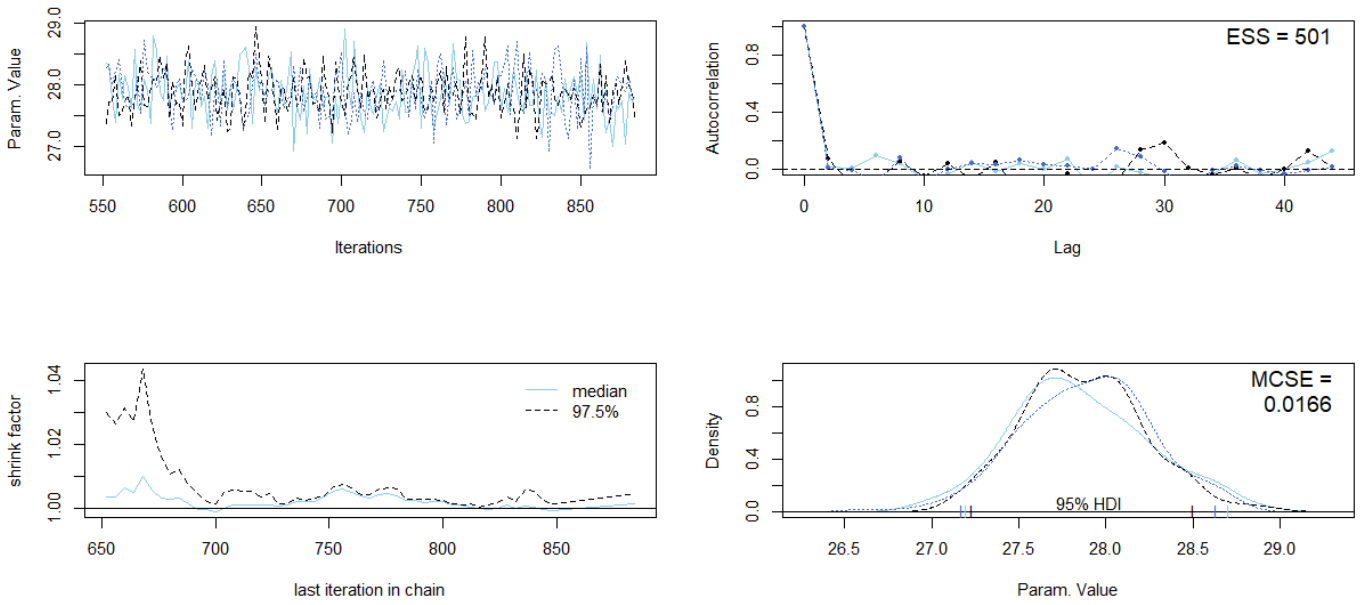

Figure 7

Same problem of Density plots and Autocorrelation is aroused.

CONCLUSION

So here, both MCMC representativeness and Accuracy are not satisfied for mean and variance, so we can't accept this posterior information.

But, we can solve the problem by changing some values of parameters.

Now, I've increase the **burn-in steps** to 200,

saved steps = 10,000 and

thinning steps = 10.

No of Chains = 3

So, I overcome the problem of MCMC Representativeness and MCMC Accuracy for both Population Mean and Population Variance.

Then, I tried different values for non-informative prior,

Case 1 : $\mu_0 = 7.5$

$\sigma_0^2 = 0.1$

$\alpha = 12960;$ $\quad \frac{\mu^2}{\sigma^2} = \frac{36^2}{0.1} = 12960,$

$\beta = 360;$ $\quad \frac{\mu}{\sigma^2} = \frac{36}{0.1} = 360,$

Case 2 : $\mu_0 = 7.5$

$\sigma_0^2 = 0.9$

$\alpha = 1440;$ $\quad \frac{\mu^2}{\sigma^2} = \frac{36^2}{0.9} = 1440,$

$\beta = 40;$ $\quad = \frac{\mu}{\sigma^2} = \frac{36}{0.9} = 40,$

Case 3 : $\mu_0 = 7.5$

$\sigma_0^2 = 1.5$

$\alpha = 864;$ $\quad \frac{\mu^2}{\sigma^2} = \frac{36^2}{1.5} = 864,$

$\beta = 24;$ $\frac{\mu}{\sigma^2} = \frac{36}{1.5} = 24,$

In every Case,

I got, nearly same estimated population mean around 6.1

And same estimated population variance around 28.

FINAL NON-INFORMATIVE PRIOR

SPECIFY THE PRIOR INFORMATION

I've put the following values to the **Gibbs-Sampling Normal-Gamma App.**

$X \sim Normal(\mu,\sigma^2)$,

$\mu \sim Normal(\mu_0,\sigma_0^2)$,

$\mu_0 = 7.5$ as given prior mean

$\sigma_0^2 = 0.5$ as low degree of belief due to Non – Informative prior

$\sigma^2 \sim Gamma(\alpha,\beta)$,

$\alpha = 2592$ as shape parameter; $\frac{\mu^2}{\sigma^2} = \frac{36^2}{0.5} = 2592,$

$\beta = 72$ as scale parameter, $\frac{\mu}{\sigma^2} = \frac{36}{0.5} = 72,$

Where, μ = given prior variance 36 = (6^2); we are given SD = 600,000,

σ = Low degree of belief cause of Non-Informative Prior.

Other MCMC Parameters,

Comparison value for the population mean = 6

Comparison value for the population variance = 27.5

No of chains = 3

Ideally, there should be only 2 to 4 chains.

No of Burn-in Steps = 200

No of Saved Steps = 10,000

No of thinning steps = 10

OUTPUT - ANALYSIS

POSTERIOR DISTRIBUTION

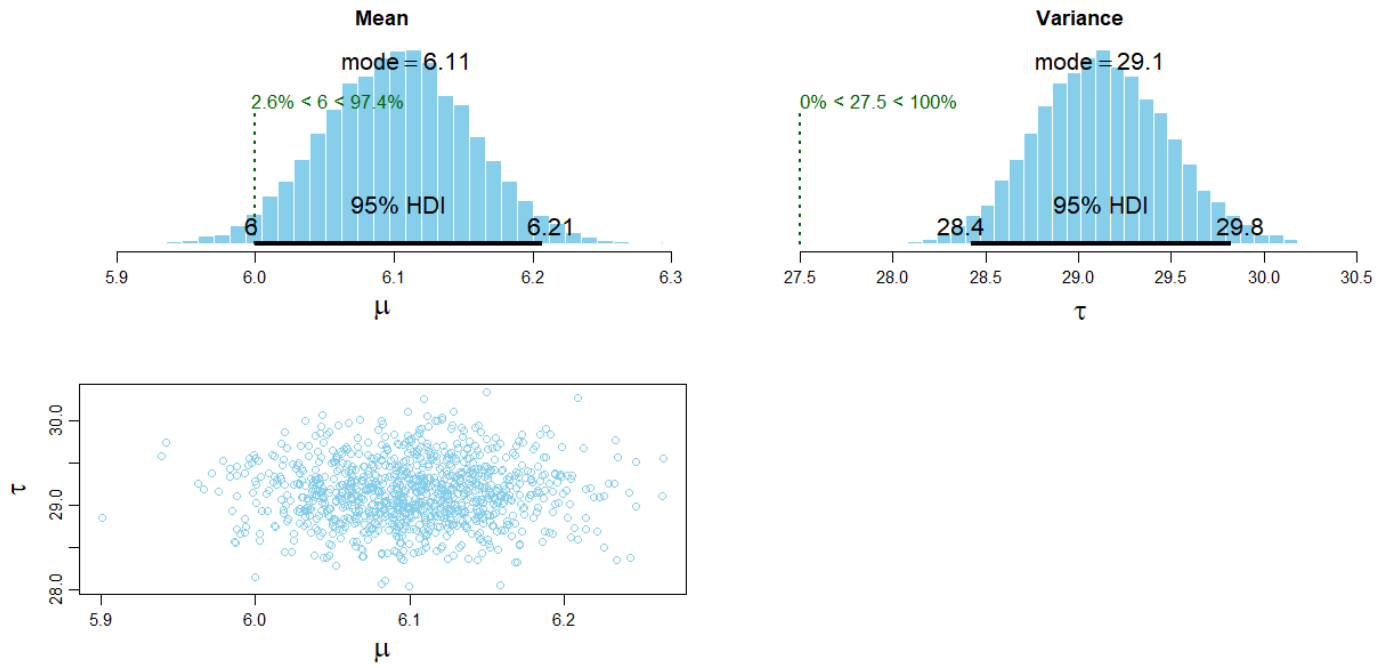

Figure 8

- As we can see from the plot that, estimated mean of population is shown as 6.11 and estimated variance of population is shown as 29.1 for 95% of HDI interval.
- These values are our Bayesian estimated mean and population of Sale Price (100K AUD) of properties of Melbourne.
- Histograms of both parameters are Symmetric due to higher number of saved steps.
- With more length of chain, we have better histograms with no gaps and gives us the idea about density of success of probability.

But, before accepting this posterior mean and variance, first we need to check MCMC Representativeness, MCMC Accuracy and MCMC Efficiency through looking into the graphs and charts of Diagnostics.

DIAGNOSTICS FOR THE MEAN, μ

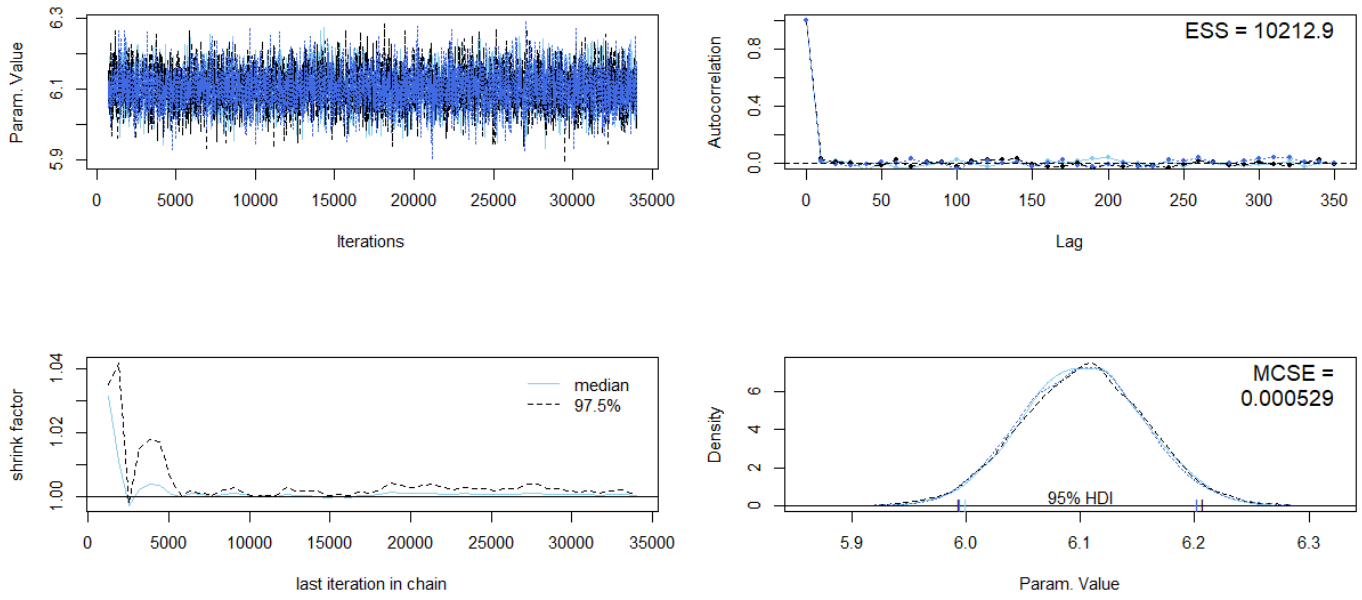

Figure 9

MCMC REPRENTATIVENESS

Trace plots give us good insights of how every chains are fluctuating around the estimated mean value. (Here, 6.1). It is crystal clear evidence that all the three chains are fluctuating around mean value 6.1, with lower and upper bound of 6.0 and 6.2 respectively.

Now, when we look at the Density plots, curves of Posterior Distribution of all the three chains are properly overlapped, even the Lower and Upper Bound of 95% HDI of all chains are somehow or nearly overlapped.

Shrink factor is 1.0 with upper bound of around 1.01. As we already know that, Shrink Factor below 1.2 is acceptable, while at the 1 is the best case.

MCMC Representativeness is checked and there down not seem any problem with MCMC Representativeness.

MCMC ACCURACY

MCMC Accuracy is mostly related with the Auto-Correlation. If there is significant auto-correlation, then these is surely a problem in the MCMC chains. We can check the Auto-Correlation while looking at the value of ESS (Estimated Sample Size). Higher the ESS value – lower the Auto-Correlation. Here, we have ESS as 10213. We are acquainted that, if ESS is greater than or equal to 10,000 – then we are on the safe side. Blue points on the chains are representing two successive points of chain given by the value of thinning step which breaks down the dependency of two successive points.

In density plots, MCSE = 0.000529 which is derived by, MCSE = $\frac{\text{SD}}{\text{ESS}}$, = $\frac{6}{10213}$ = 0.0005, which means there is 0% Monte Carlo Standard Error.

MCMC Accuracy is checked for the population mean and everything is right.

MCMC EFFICIENCY

It is directly related to the Computational power and Computational time.

DIAGNOSTICS FOR THE VARIANCE, σ^2

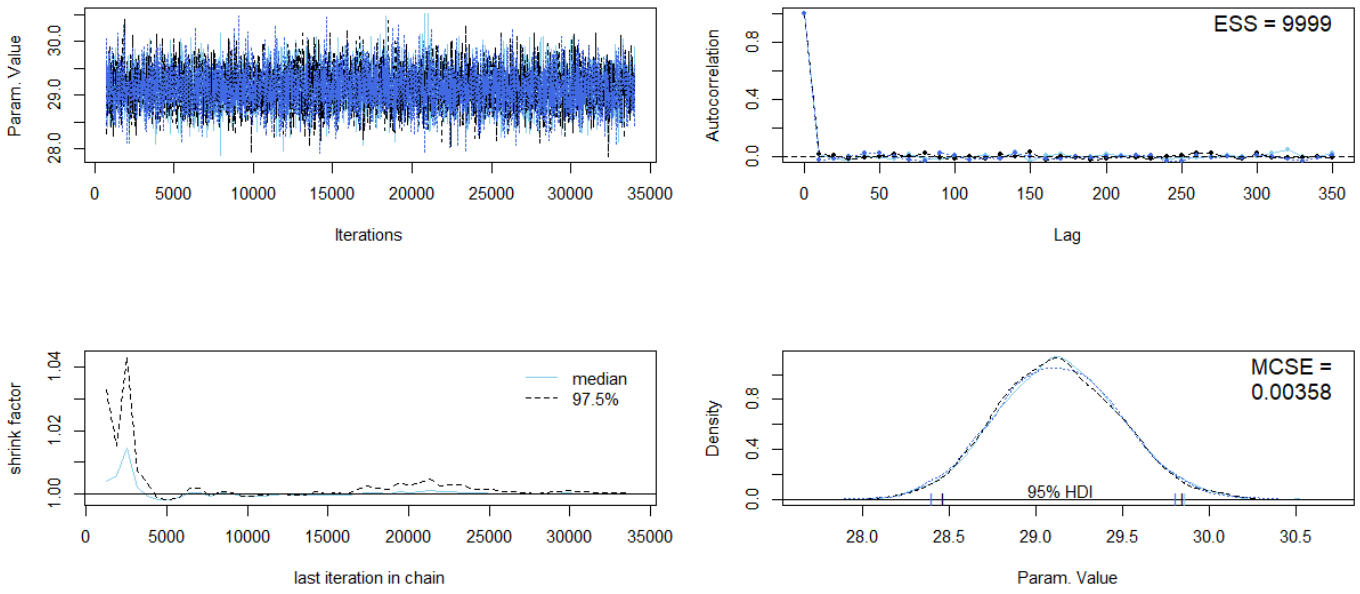

Figure 10

MCMC REPRESENTATIVENESS

We will check the same way as we did in population of mean.

Trace plots of every chains are fluctuating around the estimated variance value. (Here, 29.0). Estimated Variance value is fluctuating around 29.0 with lower and upper bound of 28.5 and 29.5.

Now, in Density plots, curves of Posterior Distribution of all the three chains are properly overlapped, with Lower and Upper Bound of 95% HDI of all chains are closer to each other.

Shrink factor is 1.0 with upper bound of around 1.001. So, Shirk factor is at the best value within acceptable range.

MCMC Representativeness is checked and everything is okay.

MCMC ACCURACY

As already discussed, Higher the ESS value – lower the Auto-Correlation. Here, we have ESS as 9999. We have been known that, if ESS is greater than or equal to 10,000 – then we are on the safe side. So , here ESS value is good. Blue points on the chains are representing two successive points of chain given by the value of thinning steps (10) which breaks down the dependency of two successive points.

In density plots, MCSE = 0.00358, which indicates there is only 0.3% Monte Carlo Standard Error.

MCMC Accuracy is checked for the population variance and there is no problem.

MCMC EFFICIENCY

It is directly related to the Computational power and Computational time.

INTERPRETATION

I have checked diagnostics of population mean and diagnostics of population variance. We must need measure of MCMC Representativeness, MCMC Accuracy as wanted, otherwise we should not go forward and infer on posterior distribution.

Here, everything is fine and we can accept this posterior distribution.

Summary of Posterior Distribution of Non-Informative prior is given below.

Posterior distributions | Summary | Diagnostics for the mean, μ | Diagnostics for the variance, τ

```
Summary table for Bayesian estimates
         Mean    Median      Mode  ESS HDImass   HDIlow
mu   6.102635  6.102597  6.104934 9999    0.95  5.99908
tau 29.130246 29.127103 29.127894 9999    0.95 28.46558
      HDIhigh CompVal PcntGtCompVal
mu   6.209467     6.0      97.24972
tau 29.840026    27.5     100.00000

Run Time
   user  system elapsed
 574.95    0.07  577.50
```

Figure 11

- mu represents estimated population mean μ,
  - As we can see, Mean-Median-Mode every value is 6.1.
  - ESS has already discussed.
  - HDI_mass = 0.95, High Density Interval of 95%
    - HDI_low (Lower bound) = 5.99, which is same as shown in Bayesian estimated posterior distribution for population mean, in the trace plots and in the density plots.
    - HDI_high (Upper bound) = 6.20, which is same as shown in Bayesian estimated posterior distribution for population mean, in the trace plots and in the density plots.
  - CompVal = 6, Comparison value for the population mean as we put in the app.
  - PcntGtCompVal = Percentage greater than comparison value for population mean.
    - 97.4 % of distribution is greater than the comparison value for population mean which is also clearly seen in the Posterior Distribution.
- tau represents estimated population mean σ,
  - As we can see, Mean-Median-Mode every value is 29.1.
  - ESS has already discussed.
  - HDI_mass = 0.95, High Density Interval of 95%
    - HDI_low (Lower bound) = 28.42, which is same as shown in Bayesian estimated posterior distribution for population mean, in the trace plots and in the density plots.
    - HDI_high (Upper bound) = 29.82, which is same as shown in Bayesian estimated posterior distribution for population mean, in the trace plots and in the density plots.
  - CompVal = 27.5, Comparison value for the population mean as we put in the app.

- PcntGtCompVal = Percentage greater than comparison value for population mean.
  - 100 % of distribution is greater than the comparison value for population mean which is also clearly seen in the Posterior Distribution.

CONCLUSION

As diagnostics of Population Mean and Population Variance is alright, we can accepts the given output of Bayesian Posterior Distribution as estimated Bayesian population mean and estimated Bayesian population variance.

Here, mean $\mu = 6.1$ and $\sigma = 29.1$

Which means for non-informative prior or low degree of belief, mean sale price(100K AUD) of properties in Melbourne is 6.1 = 610,000 AUD with standard deviation of 5.4 around the mean population.

NULL HYPOTHESES

FOR POPULATION MEAN

Test the null hypotheses that mean sale price of properties in Melbourne in 850,000 AUD.

$H_0 : \mu = 8.5$

$H_a : \mu \neq 8.5,$

Here, we will simply look at the lower and upper limit of HDI.

They are 5.99 and 6.20. $(5.99 < \mu < 6.20)$. This interval does not include 8.5.

So, we simply reject the null hypotheses as there is not strong evidence to support the null hypotheses of mean price 0f 8.5.

FOR POPULATION VARIANCE

Test the null hypotheses that standard deviance of sale price of properties in Melbourne is 300,000.

So, Standard deviation = 3,

Variance = 9

$H_0 : \sigma = 9$

$H_a : \sigma \neq 9.$

Again we will look at the HDI lower bound and upper bound.

They are 28.42 and 29.82 respectively. $(28.42 < \sigma < 29.82)$. This interval does not include 9.

So, we reject the null hypotheses as there is not strong evidence to support the null hypotheses of variance = 9.

FINAL RESULT

- Estimated Bayesian of mean of Sale price(100K AUD) = 6.1 = 610,000 AUD
- We reject null hypotheses that of mean sale price is 850,000 AUD.
- Estimated Bayesian of standard deviation is 540,000 AUD.
- We reject the null hypotheses that standard deviation of sale price of properties in Melbourne is 300,000 AUD.

As Non-Informative Prior is used, likelihood dominates the prior information.

Mean and Variance is closed to the values obtained during classical approach of Descriptive Statistics.

INFORMATIVE PRIOR

Informative prior means - very high degree of belief, low variance.

SPECIFY THE PRIOR INFORMATION

I've put the following values to the Gibbs-Sampling Normal-Gamma App.

$X \sim Normal(\mu, \sigma^2)$,

$\mu \sim Normal(\mu_0, \sigma_0^2)$,

μ_0 = 7.5 as given prior mean

σ_0^2 = 0.001 as low degree of belief due to Non – Informative prior

σ^2 ~ Gamma(α,β),

α = 1296000 as shape parameter; $\frac{\mu^2}{\sigma^2} = \frac{36^2}{0.001} = 1296000,$

β = 36000 as scale parameter, $\frac{\mu}{\sigma^2} = \frac{36}{0.001} = 36000,$

Where, **μ** = given prior variance 36 = (6^2); we are given SD = 600,000,

σ = Low degree of belief cause of Non-Informative Prior.

Other MCMC Parameters,

Comparison value for the population mean = 7

Comparison value for the population variance = 36

No of chains = 3

Ideally, there should be only 2 to 4 chains.

No of Burn-in Steps = 200

No of Saved Steps = 10,000

No of thinning steps = 10

OUTPUT - ANALYSIS

POSTERIOR DISTRIBUTION

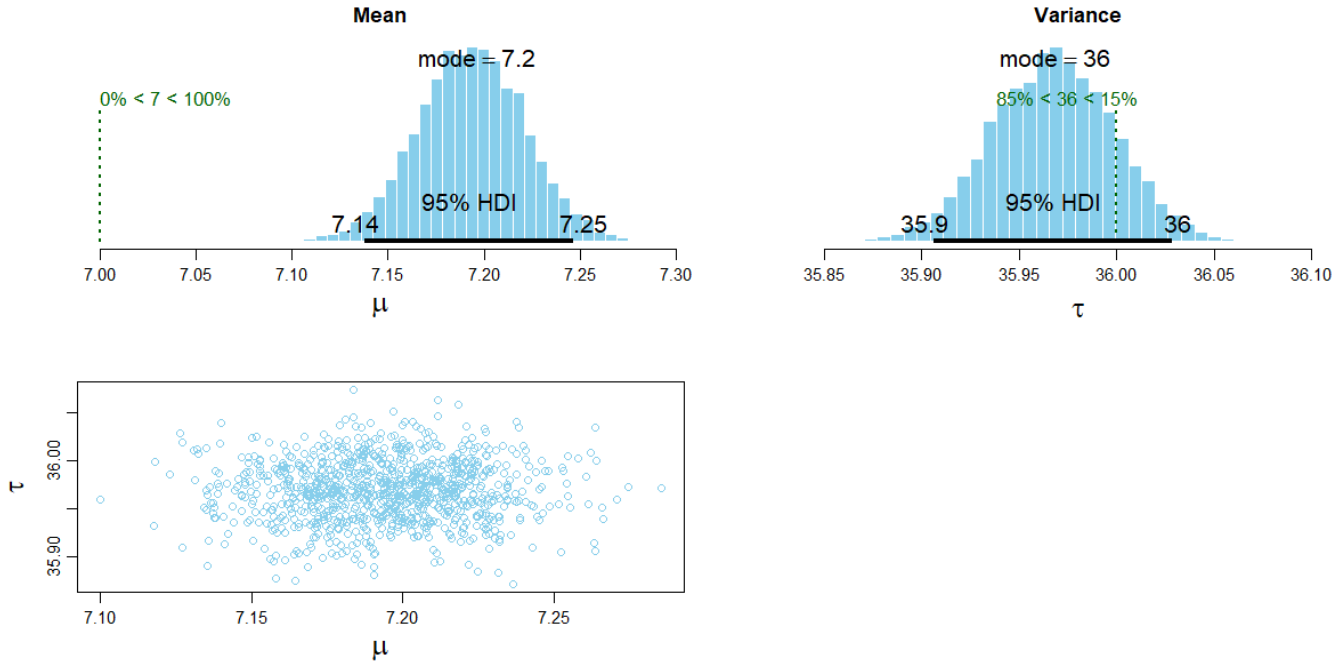

Figure 12

- As we can see from the plot that, estimated mean of population is shown as 7.2 and estimated variance of population is shown as 36 for 95% of HDI interval.
- These values are our Bayesian estimated mean and population of Sale Price (100K AUD) of properties of Melbourne.
- Histograms of both parameters are Symmetric due to higher number of saved steps.

But, before accepting this posterior mean and variance, first we need to check MCMC Representativeness, MCMC Accuracy and MCMC Efficiency through looking into the graphs and charts of Diagnostics.

DIAGNOSTICS FOR THE MEAN, μ

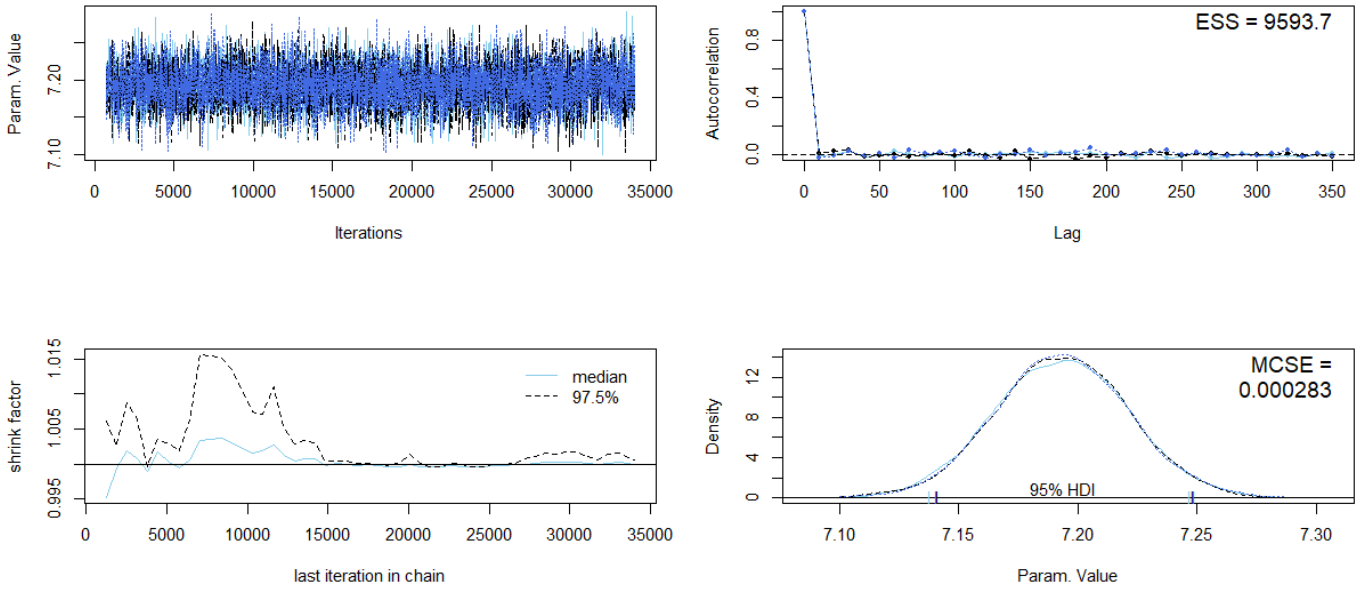

Figure 13

MCMC REPRESENTATIVENESS

Trace plots give us good insights of how every chains are fluctuating around the estimated mean value. (Here, 7.2). It is crystal clear evidence that all the three chains are fluctuating around mean value 7.2, with lower and upper bound of 7.15 and 7.25 respectively.

Now, when we look at the Density plots, curves of Posterior Distribution of all the three chains are properly overlapped, even the Lower and Upper Bound of 95% HDI of all chains are overlapped.

Shrink factor is 1.0 with upper bound of around 1.001. As we already know that, Shrink Factor below 1.2 is acceptable, while at the 1 is the best case.

MCMC Representativeness is checked and there down not seem any problem with MCMC Representativeness.

MCMC ACCURACY

MCMC Accuracy is mostly related with the Auto-Correlation. If there is significant auto-correlation, then these is surely a problem in the MCMC chains. We can check the Auto-Correlation while looking at the value of ESS (Estimated Sample Size). Higher the ESS value – lower the Auto-Correlation. Here, we have ESS as 9593.7. ESS value is high, which is extremely good scenario. Blue points on the chains are representing two successive points of chain given by the value of thinning step which breaks down the dependency of two successive points.

In density plots, MCSE = 0.000283, which means there is 0% Monte Carlo Standard Error.

MCMC Accuracy is checked for the population mean and everything is right.

MCMC EFFICIENCY

It is directly related to the Computational power and Computational time.

DIAGNOSTICS FOR THE VARIANCE, σ^2

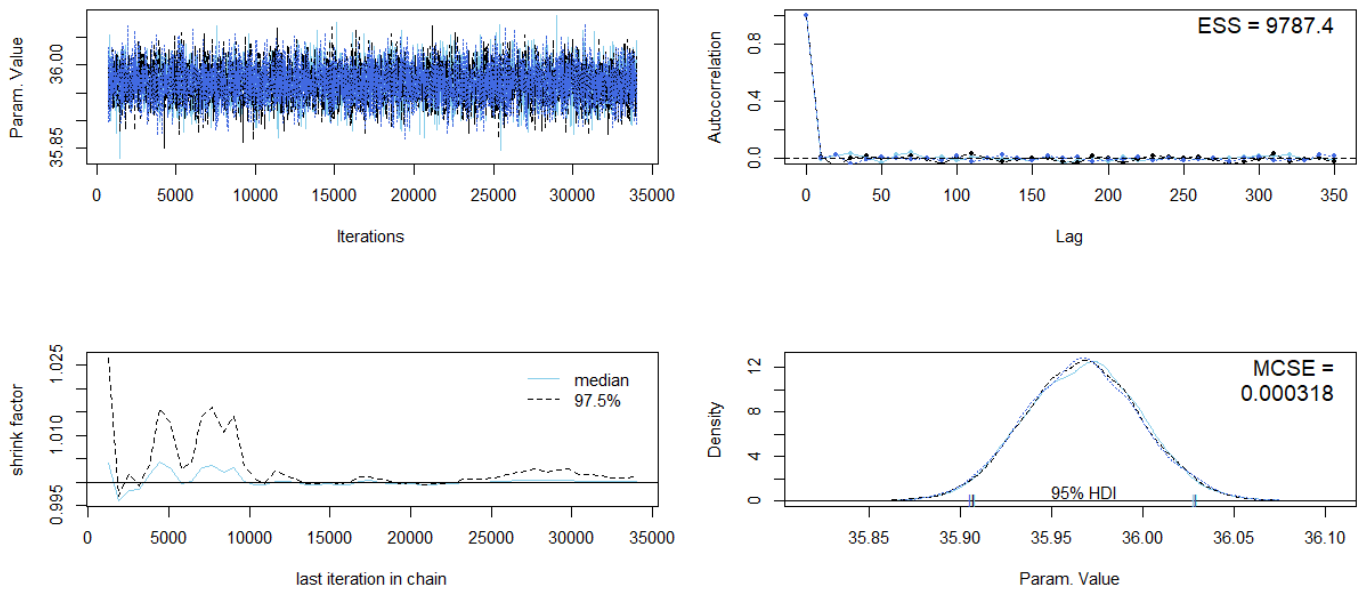

Figure 14

MCMC REPRESENTATIVENESS

We will check the same way as we did in population of mean.

Trace plots of every chains are fluctuating around the estimated variance value. (Here, 36.0). Estimated Variance value is fluctuating around 36.0 with lower and upper bound of 35.9 and 36.01 respectively.

Now, in Density plots, curves of Posterior Distribution of all the three chains are properly overlapped, with Lower and Upper Bound of 95% HDI of all chains are overlapped or extremely closed to each other.

Shrink factor is 1.0 with upper bound of around 1.001. So, Shirk factor is at the best value within acceptable range.

MCMC Representativeness is checked and everything is okay.

MCMC ACCURACY

As already discussed, Higher the ESS value – lower the Auto-Correlation. Here, we have ESS as 9787.4. So, here ESS value is good. Blue points on the chains are representing two successive points of chain given by the value of thinning steps (10) which breaks down the dependency of two successive points.

In density plots, MCSE = 0.00318, which indicates there is only 0% Monte Carlo Standard Error.

MCMC Accuracy is checked for the population variance and there is no problem.

MCMC EFFICIENCY

It is directly related to the Computational power and Computational time.

INTERPRETATION

I have checked diagnostics of population mean and diagnostics of population variance. We must need measure of MCMC Representativeness, MCMC Accuracy as wanted, otherwise we should not go forward and infer on posterior distribution.

Here, everything is fine and we can accept this posterior distribution.

Summary of Posterior Distribution of Non-Informative prior is given below.

Posterior distributions | Summary | Diagnostics for the mean, μ | Diagnostics for the variance, τ

```
Summary table for Bayesian estimates
         Mean    Median      Mode    ESS HDImass    HDIlow   HDIhigh CompVal
mu   7.193251  7.193232  7.196263 9421.7    0.95  7.138023  7.246805       7
tau 35.967437 35.967655 35.968387 9999.0    0.95 35.906420 36.028399      36
    PcntGtCompVal
mu       100.0000
tau       15.0115

Run Time
   user  system elapsed
 661.02    0.08  665.59
```

Figure 15

- mu represents estimated population mean μ,
  - As we can see, Mean-Median-Mode every value is 7.2
  - ESS has already discussed.
  - HDI_mass = 0.95, High Density Interval of 95%
    - HDI_low (Lower bound) = 7.14, which is same as shown in Bayesian estimated posterior distribution for population mean, in the trace plots and in the density plots.
    - HDI_high (Upper bound) = 7.25, which is same as shown in Bayesian estimated posterior distribution for population mean, in the trace plots and in the density plots.
  - CompVal = 7, Comparison value for the population mean as we put in the app.
  - PcntGtCompVal = Percentage greater than comparison value for population mean.
    - 100 % of distribution is greater than the comparison value for population mean which is also clearly seen in the Posterior Distribution.
- tau represents estimated population mean σ,
  - As we can see, Mean-Median-Mode every value is 35.97
  - ESS has already discussed.
  - HDI_mass = 0.95, High Density Interval of 95%
    - HDI_low (Lower bound) = 35.90, which is same as shown in Bayesian estimated posterior distribution for population mean, in the trace plots and in the density plots.
    - HDI_high (Upper bound) = 36.03, which is same as shown in Bayesian estimated posterior distribution for population mean, in the trace plots and in the density plots.
  - CompVal = 36, Comparison value for the population mean as we put in the app.
  - PcntGtCompVal = Percentage greater than comparison value for population mean.
    - 15 % of distribution is greater than the comparison value for population mean which is also clearly seen in the Posterior Distribution.

CONCLUSION

As diagnostics of Population Mean and Population Variance is alright, we can accepts the given output of Bayesian Posterior Distribution as estimated Bayesian population mean and estimated Bayesian population variance.

Here, mean μ = 7.2 and σ = 36

Which means for non-informative prior or low degree of belief, mean sale price(100K AUD) of properties in Melbourne is 7.2 = 720,000 AUD with standard deviation of 6 around the mean population.

NULL HYPOTHESE

FOR POPULATION MEAN

Test the null hypotheses that mean sale price of properties in Melbourne in 850,000 AUD.

$H_0 : \mu = 8.5$

$H_a : \mu \neq 8.5,$

Here, we will simply look at the lower and upper limit of HDI.

They are 7.14 and 7.25 ($7.14 < \mu < 7.25$). This interval does not include 8.5.

So, we simply reject the null hypotheses as there is not strong evidence to support the null hypotheses of mean price 0f 8.5.

FOR POPULATION VARIANCE

Test the null hypotheses that standard deviance of sale price of properties in Melbourne is 300,000.

So, Standard deviation = 3,

Variance = 9

$H_0 : \sigma = 9$

$H_a : \sigma \neq 9.$

Again we will look at the HDI lower bound and upper bound.

They are 35.9 and 36.03 respectively. ($35.9 < \sigma < 36.03$). This interval does not include 9.

So, we reject the null hypotheses as there is not strong evidence to support the null hypotheses of variance = 9.

FINAL RESULT

- Estimated Bayesian of mean of Sale price(100K AUD) = 7.2 = 720,000 AUD
- We reject null hypotheses that of mean sale price is 850,000 AUD.
- Estimated Bayesian of standard deviation is 599,000 AUD.
- We reject the null hypotheses that standard deviation of sale price of properties in Melbourne is 300,000 AUD.

As Informative Prior is used, prior distribution dominates the likelihood.

As estimated mean is 720,000 AUD, which is close to 750,000 of prior information than 610,000 AUD that we obtained during the classical approach of descriptive statistics.

As estimated standard deviation is 599,000 AUD which is again close to prior information than what we obtained during classical descriptive statistics.

APPENDICES

I've taken data for this report from the assignment 1 itself. Rename the column "Sale Price" as "y", which is essential to obtain output from the Gibbs-Sampling Normal-Gamma Application.

Some important terminology for this report are Burn-in, Shrink Factor, Auto-Correlation, ESS (Estimated Sample Size), Lag and Thinning Steps. Which are described in Some Important Terminology Section.

I've entirely create this report from Class notes, Lectures and Recordings only.

In some parts, I've used term "Target Distribution" which is "Posterior Distribution".

Normal Distribution is also known as Gaussian Distribution.

REFERENCES

- Class Lectures – Notes
- Class presentations
- Class Recordings