ST4060 – Statistical Methods for Machine Learning I
ST6015 – Computer Analytical Techniques for Actuarial Applications
ST6040 – Machine Learning and Statistical Analytics I

Continuous Assessment 1 2020-21

Name: ADITYA VIKRAM DASH
Student ID: 120224162
Programme (FMAS4, RAS4, MScDSA, etc.): MScDSA

Please provide your answers to each question below, and provide any relevant R code with your answers to each question. Submit a pdf version of this document as your unique submission document on Canvas.

Remember to save your work regularly.

Your answer to Question 1:

(a) 1

(b) h

(c) No

(d) h1; Less bias

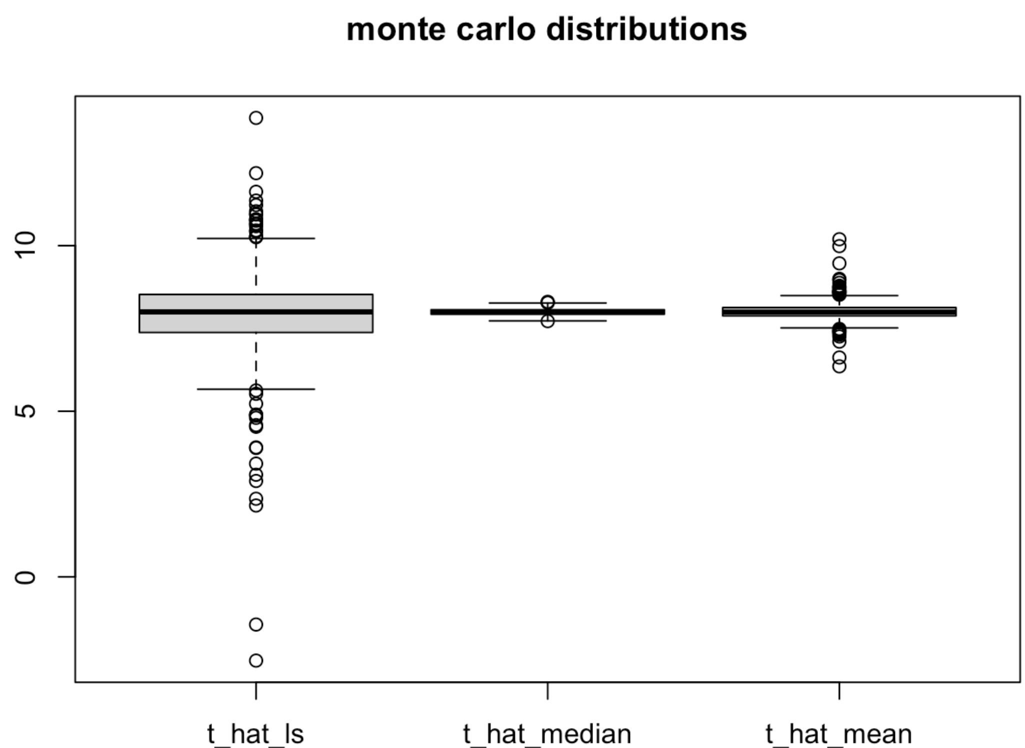
Your answers to Question 2:

(a) 8.004457, 7.99634, 8.007675

(b) 0.007042658, 0.003065645, 0.008026896

(c) We should go for the estimator which has low bias and low standard error. se_t_med and bias_est1 have the least values, so we should go for the second estimator.

(d)



Relevant R code for Question 2:

```
print("Estimated Values are as : ")

#Estimate_1
exp_t_ls = mean(est_t_hat)
exp_t_ls

#Estimate_2
exp_t_med = mean(e_t_med)
exp_t_med

#Estimate_3
exp_t_mu = mean(estimates_t_hat_mean)
exp_t_mu
```

```
print("Standard Errors are as :")

#SE_Estimate_1
se_t_ls = sd(est_t_hat)/sqrt(M)
se_t_ls

#SE_Estimate_2
se_t_med = sd(e_t_med)/sqrt(M)
se_t_med #SE_Estimate 2

#SE_Estimate_3
se_t_mu = sd(estimates_t_hat_mean)/sqrt(M)
se_t_mu #SE_Estimate 3
```

```
#Boxplot
data <- data.frame(t_hat_ls=est_t_hat, t_hat_median=e_t_med, t_hat_mean=estimates_t_hat_mean)
boxplot(data,main="monte carlo distributions")
```

<u>Your answers to Question 3:</u>

(a) – 5.419382

(b) 1.228608e-07, 4.620393e-14

(c) (-6.572610, -3.773033)

(d) (-6.440310, -4.248634)

On comparing the Bootstrap Confidence Interval with the traditional 95% confidence interval, we see that confidence interval original (w/o bootstrapping) assumes normality, whereas bootstrap confidence interval doesn't. So we should go with bootstrapping.

Relevant R code for Question 3:

```
#Expected Value of Slope Coefficient
bts_exp_slope = mean(est_slopes)
```

```
#Naive
ci_p_value = quantile(est_p_values,c(0.025,0.975))
```

```
#Bootstrap 95% CI
bts_ci = c(2*(true_estimate)-quantile(est_slopes,c(0.025,0.975))[2], (2*true_estimate)-quantile(est_slopes,c(0.025,0.975))[1])
```

```
#W/o Bootstrapping
original_slope_ci = c(coef(model)[2]-1.96*(summary(model)$coefficients[2,2]),coef(model)[2]+1.96*(summary(model)$coefficients[2,2]))
```