0413作業

計財所 碩二 108071601 賴冠維

2021/4/12

## Problem 1

### (a)

從抽樣計算積分式，所得平均數及變異數如下：

## [1] 0.52533583 0.05991423

若則與原式相同，服從,計算平均數及變異數如下：

## [1] "Mean of the values of simulation Integration: 0.524855942320723"

## [1] "Variance of the values of simulation Integration: 0.0601152584276106"

### (b)

若，則變成服從分配並且積分分配值域為，計算其平均數及變異數如下：

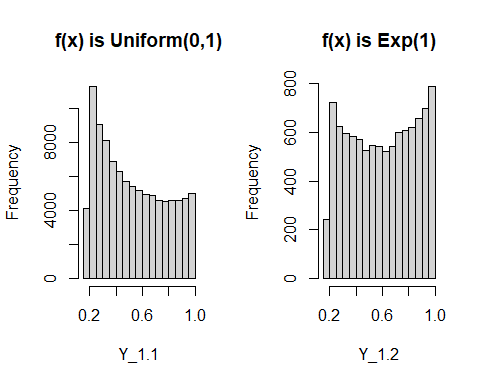
## [1] "Mean of the values of simulation Integration: 0.59917234298382"

## [1] "Variance of the values of simulation Integration: 0.0612878170271518

### (c)

#### (a),(b)小題有何差別？

* (a)小題為計算均勻分配的積分式，而(b)小題則是計算指數分配的積分式,可發現指數分配在機率分布上沒有均勻分配來的平穩，可能來自我們限制x的值域落在之間，原指數分配的值域為，僅擷取指數分配中的一小段,導致偏誤較大。



## Problem 2

### 建立模擬樣本

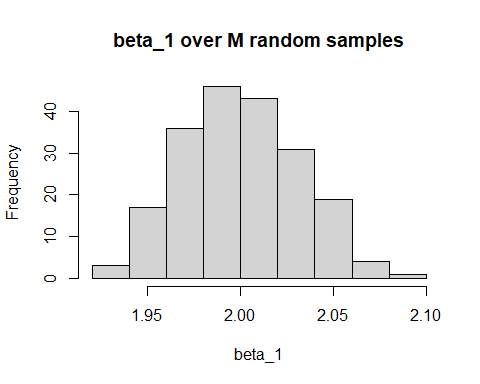
先根據模擬數據

### (a)

可由下圖發現在大樣本下其估計出的Variance極低約為0.001，從直方圖來看呈現鐘型曲線服從常態分配。

## [1] "Mean of beta\_1: 2.00086840409782"

## [1] "Variance of beta\_1: 0.00096465667958181"



### (b)

可發現在估計時，其值隨著每次模擬樣本的更動而有所不同，在這200個的估計值中，全距大約0.156上下，而Boostrap後的平均數卻可讓的估計值穩健地落在2。

## [1] "Maximum in the estimator of beta\_1: 2.08548410994247"

## [1] "Minimum in the estimator of beta\_1: 1.92855244270411"

## [1] "The range of the estimator of beta\_1: 0.156931667238352"

### (c)

以計算其變異數後平均，如下：

## [1] "Mean of the Asymptotic Variance of beta\_1: 0.0040090830103712"

計算，算出200組的變異數，如下：

## [1] "Mean of the Empirical Variance of beta\_1: 0.00096465667958181"

可發現兩者在樣本數很大的情況下，變異數都非常小。

## [1] "Mean of the Asymptotic Variance of beta\_1: 0.0040090830103712"

## [1] "Mean of the Empirical Variance of beta\_1: 0.00096465667958181"

### (d)

第一種為的，先從200組資料中，對每組的500筆資料重抽樣，可得新的200組樣本，對這200組計算，計算其變異數，如下：

## [1] "Mean of the estimator of beta\_1 by Random x boostrap: 2.00753959826958"

## [1] "Variance of the estimator of beta\_1 by Random x boostrap: : 0.00353469030345969"

第二種為的，先將原200組資料配適可得殘差，接著將殘差做200組重抽樣後，取出對應的X與其配適值，再將殘差加上配適值後得到新的，以計算，接著計算其變異數，如下：

## [1] "Mean of the estimator of beta\_1 by Fixed x boostrap: 1.99604171567323"

## [1] "Variance of the estimator of beta\_1 by Fixed x boostrap: 0.00318864235815397"

# (e)

為在中多乘上一個機率分配,並且，可得的為，因此將乘上，計算如下：  
**(a)小題加上擾動項**

## [1] "Variance of Perturbation\_beta\_1 (a): 0.0268585508738842"

### (d)小題加上擾動項

## [1] "Variance of Perturbation\_beta\_1 (d) Random x: 0.0760376069539746"

## [1] "Variance of Perturbation\_beta\_1 (d) Fixed x: 0.0630605144545954"

觀察各題結果，發現使用會使得變異數明顯大於其他方法所求，而其他方法在樣本數足夠大下，其變異數皆非常小。

## [1] "Variance of beta\_1 in (a): 0.00096465667958181"

## [1] "Mean of the Asymptotic Variance of beta\_1 in (c): 0.0040090830103712"

## [1] "Mean of the Empirical Variance of beta\_1 in (c): 0.00096465667958181"

## [1] "Variance of the estimator of beta\_1 by Random x boostrap in (d): 0.0035346903034"

## [1] "Variance of the estimator of beta\_1 by Fixed x boostrap in (d): 0.00318864235815"

## [1] "Variance of Perturbation\_beta\_1 for (a) in (e): 0.0268585508738842"

## [1] "Variance of Perturbation\_beta\_1 for (d) Random x in (e): 0.0760376069539746"

## [1] "Variance of Perturbation\_beta\_1 for (d) Fixed x in (e): 0.0630605144545954"

附錄(程式碼)

library(tidyverse)

## Problem 1

### (a)

set.seed(123)

X = runif(100000,0,1)

Y = (exp(-X)/(1+X^2))

c( mean(Y), var(Y) )

w <- function(x) (exp(-X)/(1+X^2))/dunif(x, 0, 1)

f <- function(x) dunif(x,0,1)

X = runif(100000,0,1)

Y\_1.1 = w(X)\*f(X)

paste0("Mean of the values of simulation Integration: ",mean(Y\_1.1))

paste0("Variance of the values of simulation Integration: ",var(Y\_1.1))

### (b)

w <- function(x) 1/(1+x^2)

f <- function(x) dexp(x,rate = 1)

X = matrix(NA,nrow = 10000)

i = 1

while (i< 10001) {

e = rexp(1,rate = 1)

if (e<1) {

X[i] = e

i = i+1

}

}

Y\_1.2 = w(X)\*f(X)

paste0("Mean of the values of simulation Integration: ",mean(Y\_1.2))

paste0("Variance of the values of simulation Integration: ",var(Y\_1.2))

par(mfrow=c(1,2))

hist(Y\_1.1,main = "f(x) is Uniform(0,1)")

hist(Y\_1.2,main = "f(x) is Exp(1)")

## Problem 2

set.seed(1234)

Simulation\_Data = lapply(1:200, function(a){

X = rnorm(n = 500,mean = 0,sd = sqrt(2))

Y = 1+2\*X+rnorm(n = 500,mean = 0,sd = 1)

list(X,Y)

})

names(Simulation\_Data) = paste0("Trial\_",1:200)

### (a)

beta\_1 = sapply(1:200, function(a){

trial = Simulation\_Data[[a]]

trial\_lm = lm(trial[[2]]~trial[[1]])

beta\_1 = trial\_lm$coefficients[[2]]

beta\_1

})

paste0("Mean of beta\_1: ",mean(beta\_1))

paste0("Variance of beta\_1: ",var(beta\_1))

hist(beta\_1,main = "beta\_1 over M random samples")

paste0("Maximum in the estimator of beta\_1: ",max(beta\_1))

paste0("Minimum in the estimator of beta\_1: ",min(beta\_1))

paste0("The range of the estimator of beta\_1: ",max(beta\_1)-min(beta\_1))

### (c)

Var\_MLE = sapply(1:200, function(a){

x = Simulation\_Data[[a]][[1]]

y = Simulation\_Data[[a]][[2]]

(sum(y\*x)/sum(x\*x))/length(y)

})

paste0("Mean of the Asymptotic Variance of beta\_1: ",mean(Var\_MLE))

a = sapply(1:200, function(a){

LM = lm(Simulation\_Data[[a]][[2]]~Simulation\_Data[[a]][[1]])

beta\_1 = LM$coefficients[[2]]

beta\_1

})

paste0("Mean of the Empirical Variance of beta\_1: ",var(a))

paste0("Mean of the Asymptotic Variance of beta\_1: ",mean(Var\_MLE))

paste0("Mean of the Empirical Variance of beta\_1: ",var(a))

### (d)

set.seed(12345)

Bst\_Data = lapply(1:200, function(a){

index = sample(1:500,200,replace = T)

d = cbind(X = Simulation\_Data[[a]][[1]][index],

Y=Simulation\_Data[[a]][[2]][index]) %>%

as.data.frame()

})

Obs\_Bst\_Var = sapply(1:200, function(a){

L = lm(Bst\_Data[[a]]$Y~Bst\_Data[[a]]$X)

var = L$coefficients[[2]]

var

})

paste0("Mean of the estimator of beta\_1 by Random x boostrap: ",mean(Obs\_Bst\_Var))

paste0("Variance of the estimator of beta\_1 by Random x boostrap: : ",var(Obs\_Bst\_Var))

set.seed(12345)

LM = lapply(1:200, function(a){

Origin\_LM = lm(Simulation\_Data[[a]][[2]]~Simulation\_Data[[a]][[1]])

list(res = Origin\_LM$residuals,

fit = Origin\_LM$fitted.values,

x = Origin\_LM$model$`Simulation\_Data[[a]][[1]]`)

})

Res\_Bst\_Data = lapply(1:200, function(a){

BS\_index = sample(1:500,200,replace = T)

index = sample(1:500,200,replace = F)

res = LM[[a]]$res[BS\_index]

x = LM[[a]]$x[index]

Y\_1 = LM[[a]]$fit[index]+res

cbind(x,Y\_1) %>% as.data.frame()

})

Res\_Bst\_Var = sapply(1:200, function(a){

L = lm(Res\_Bst\_Data[[a]]$Y\_1~Res\_Bst\_Data[[a]]$x)

var = L$coefficients[[2]]

var

})

paste0("Mean of the estimator of beta\_1 by Fixed x boostrap: ",mean(Res\_Bst\_Var))

paste0("Variance of the estimator of beta\_1 by Fixed x boostrap: ",var(Res\_Bst\_Var))

# (e)

set.seed(12345)

per\_beta\_1 = sapply(1:200, function(a){

trial = Simulation\_Data[[a]]

x = rexp(n = 500,rate = 1)

per\_x = trial[[1]]

per\_y = trial[[2]]\*x

trial\_lm = lm(per\_y~per\_x)

beta\_1 = trial\_lm$coefficients[[2]]

beta\_1

})

paste0("Variance of Perturbation\_beta\_1 (a): ",var(per\_beta\_1))

set.seed(12345)

Per\_Bst\_Data = lapply(1:200, function(a){

index = sample(1:500,200,replace = T)

d = cbind(X = Simulation\_Data[[a]][[1]][index],

Y=Simulation\_Data[[a]][[2]][index]\*rexp(n = 200,rate = 1)) %>%

as.data.frame()

})

Per\_Obs\_Bst\_Var = sapply(1:200, function(a){

L = lm(Per\_Bst\_Data[[a]]$Y~Per\_Bst\_Data[[a]]$X)

var = L$coefficients[[2]]

var

})

paste0("Variance of Perturbation\_beta\_1 (d) Random x: ",var(Per\_Obs\_Bst\_Var))

set.seed(12345)

LM = lapply(1:200, function(a){

Origin\_LM = lm(Simulation\_Data[[a]][[2]]~Simulation\_Data[[a]][[1]])

list(res = Origin\_LM$residuals,

fit = Origin\_LM$fitted.values,

x = Origin\_LM$model$`Simulation\_Data[[a]][[1]]`)

})

Per\_Res\_Bst\_Data = lapply(1:200, function(a){

BS\_index = sample(1:500,200,replace = T)

index = sample(1:500,200,replace = F)

res = LM[[a]]$res[BS\_index]

x = LM[[a]]$x[index]

Y\_1 = (LM[[a]]$fit[index]+res)\*rexp(n = 200,rate = 1)

cbind(x,Y\_1) %>% as.data.frame()

})

Per\_Res\_Bst\_Var = sapply(1:200, function(a){

L = lm(Per\_Res\_Bst\_Data[[a]]$Y\_1~Per\_Res\_Bst\_Data[[a]]$x)

var = L$coefficients[[2]]

var

})

paste0("Variance of Perturbation\_beta\_1 (d) Fixed x: ",var(Per\_Res\_Bst\_Var))

paste0("Variance of beta\_1 in (a): ",var(beta\_1))

paste0("Mean of the Asymptotic Variance of beta\_1 in (c): ",mean(Var\_MLE))

paste0("Mean of the Empirical Variance of beta\_1 in (c): ",var(a))

paste0("Variance of the estimator of beta\_1 by Random x boostrap in (d): ",var(Obs\_Bst\_Var))

paste0("Variance of the estimator of beta\_1 by Fixed x boostrap in (d): ",var(Res\_Bst\_Var))

paste0("Variance of Perturbation\_beta\_1 for (a) in (e): ",var(per\_beta\_1))

paste0("Variance of Perturbation\_beta\_1 for (d) Random x in (e): ",var(Per\_Obs\_Bst\_Var))

paste0("Variance of Perturbation\_beta\_1 for (d) Fixed x in (e): ",var(Per\_Res\_Bst\_Var))