R07849035 鄭凱元 homework 3

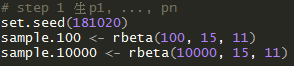
Exercise 1-1

1. Use Monte Carlo Integration to estimate probability of favoring success?

請嘗試不同的MC抽樣個數(100, 10000)並詳細解釋每個步驟

Step 1: 生

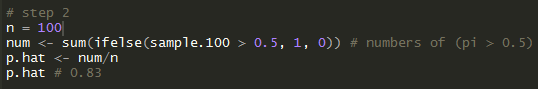
使用R中的rbeta()指令生出100個或10000個之間且服從Beta(15, 11)的隨機樣本，並使用set.seed()將這次抽出的隨機樣本存起來。



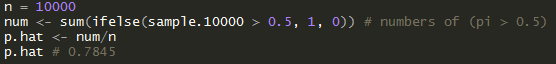
Step 2:

將生出的樣本帶進Indicator function內，若樣本值大於0.5為則輸出1、若樣本值小於0.5則輸出0；再將所有輸出結果加總後除以總個數。

在樣本數為100的時候，得到。



在樣本數為10000的時候，得到。



Step 3:

計算之變異數

在樣本數為100的時候，得到。



在樣本數為10000的時候，得到。



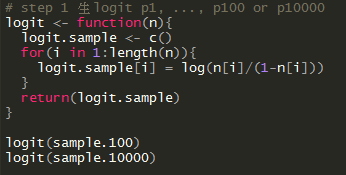
Exercise 1-2

1. Use Monte Carlo Integration to estimate ?

請嘗試不同的MC抽樣個數(100, 10000)並詳細解釋每個步驟

Step 1: 生並轉換成

這裡延續使用上一題所隨機生出之樣本，將每個樣本轉換成logit的形式。



Step 2:

樣本數為100時，可得為0.3052098；

樣本數為10000時，可得為0.3242664。



Step 3:

樣本數為100時，可得為0.001249733；

樣本數為10000時，可得為。



Exercise 2

Use Important sampling to complete previous two problems

Uniform(0, 1) as the important function and and

完成後請跟前面Monte Carlo Integration比較

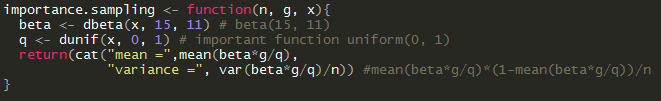
Step 1: 生

使用R中的runif()指令生出100個或10000個之間且服從Uniform(0, 1)的隨機樣本，並將樣本存進指定的變數名稱中。



Step 2&3

此處我寫了一個function，function中包含需要用到的posterior distribution Beta(15, 11)與importance function Uniform(0, 1)。而輸入樣本數n、函數與posterior distribution和importance function的抽樣母體分配x後就會輸出的期望值與變異數。



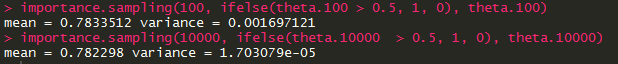
將Exercise 1-1之函數輸入之後，因為變異數的部分有些問題，所以將變異數公式的部分改成第一題題目所要求的：

當樣本數為100時，

可得 與 ；

當樣本數為10000時，

可得 與 。



將Exercise 1-2之函數輸入之後，將會得到以下結果：

當樣本數為100時，

可得 與 ；

樣本數為10000時，

可得 與 。



Monte Carlo Integration與Important sampling的比較：

：

|  |  |  |
| --- | --- | --- |
| sample = 100 | Monte Carlo | Important sampling |
|  | 0.83 | 0.7833512 |
|  | 0.001411 | 0.001697121 |

|  |  |  |
| --- | --- | --- |
| sample = 10000 | Monte Carlo | Important sampling |
|  | 0.7845 | 0.782298 |
|  |  |  |

：

|  |  |  |
| --- | --- | --- |
| sample = 100 | Monte Carlo | Important sampling |
|  | 0.3052098 | 0.2439954 |
|  | 0.001249733 | 0.003980534 |

|  |  |  |
| --- | --- | --- |
| sample = 10000 | Monte Carlo | Important sampling |
|  | 0.3242664 | 0.3315126 |
|  |  |  |

由上面所作之練習，可以發現Monte Carlo Integration的樣本是由Beta(15, 11)中所抽樣而得，而Important sampling的樣本則是由Uniform(0, 1)中所隨機抽取，故跑出來的結果不會完全一致，但理論上也不會差太多。

將跑出來的數字做個表格整理，可以看出當樣本數為100時，變異數都會相較樣本數為10000時來的大；兩種方法做出來的期望值也是在樣本數為100時差的比較多，不過我認為是變異數造成的結果。

Importance sampling是當posterior distribution為不常見的分配時同乘除一個較熟悉的分配再做運算，所以我認為兩種方法並無優劣，其差異是在使用的時機，當然如果能直接用Monte Carlo Integration的話還是直接使用Monte Carlo Integration。