

# 專題討論2





# 問題一： MCMC的Acceptance rate



# 問題一：來自rejection method想法

- 步驟一：

從一個容易取得的distribution  $g(X)$ 中，隨機取一個位置座標 $z$  (在任意位置上  $m \times g(x) > f(x)$ )

- 步驟二：

計算ratio  $R = \frac{f(z)}{m \times g(z)}$  (這個 $R$ 可想為accept的機率)

- 步驟三：

從uniform(0,1)中隨機取樣一個 $u$ ，比較 $R$ 與：

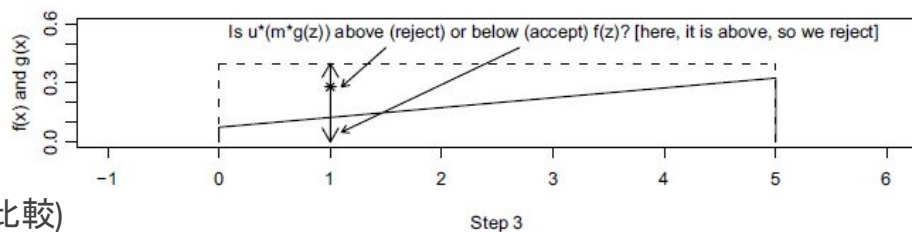
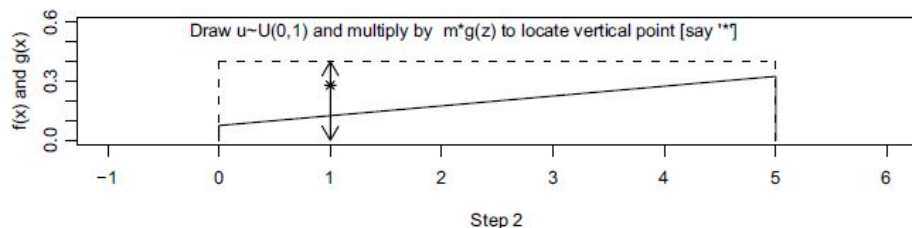
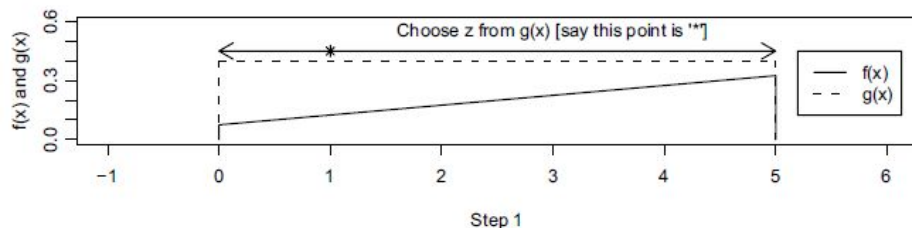
If  $R > u$ , accept  $z$

Otherwise, reject  $z$  & return to step 1.

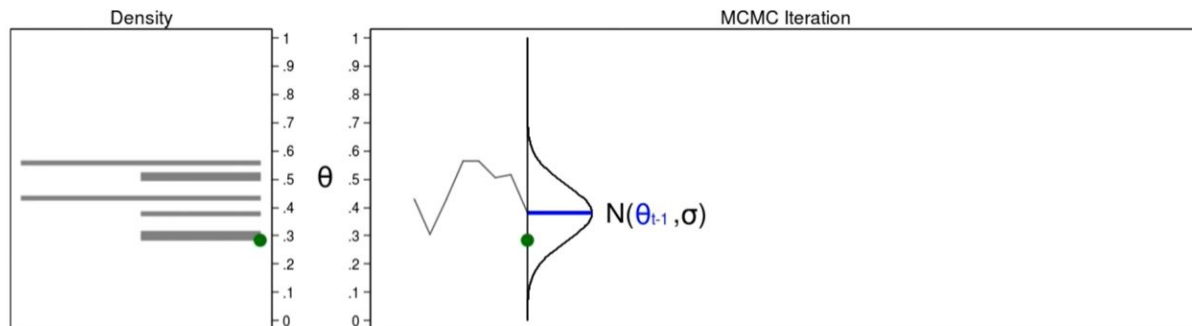
**Note:**

以上步驟或可以想像為： $f(z) \leq m \times g(z) \times u$   
(右側乘上 $u$ 為從 $m \times g(z)$ 的

這個高度中隨機取樣，並與左側目標分布在此位置上的density比較)



# 問題一：MCMC的Acceptance rate



$$\text{Step 1: } r(\theta_{\text{new}}, \theta_{t-1}) = \frac{\text{Posterior}(\theta_{\text{new}})}{\text{Posterior}(\theta_{t-1})} = \frac{\text{Beta}(1,1,0.286) \times \text{Binomial}(10,4,0.286)}{\text{Beta}(1,1,0.380) \times \text{Binomial}(10,4,0.380)} = 0.747$$

$$\text{Step 2: Acceptance probability } \alpha(\theta_{\text{new}}, \theta_{t-1}) = \min\{r(\theta_{\text{new}}, \theta_{t-1}), 1\} = \min\{0.747, 1\} = 0.747$$

$$\text{Step 3: Draw } u \sim \text{Uniform}(0,1) = 0.094$$

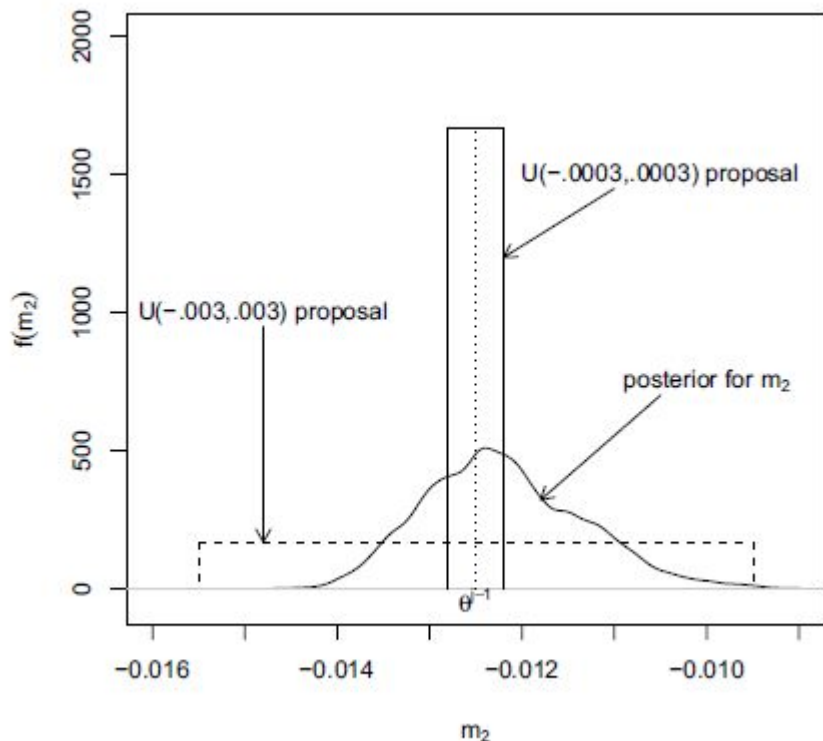
$$\text{Step 4: If } u < \alpha(\theta_{\text{new}}, \theta_{t-1}) \rightarrow \text{If } 0.094 < 0.747 \quad \text{Then } \theta_t = \theta_{\text{new}} = 0.286 \\ \text{Otherwise } \theta_t = \theta_{t-1} = 0.380$$

Step 1算前後model的ratio就像是 Bayes factor(見下面連結)中想要比較model 1 or model 2誰fit較好，如果model 1 fit好，這個ratio就會大於1，我們一定接受，所以 acceptance probability=1

如果ratio小於1，則就像rejection method，這個ratio可以想為accept的機率，如果大於uniform(0,1)隨機抽樣的結果u，則accept這個新的位置下的model (例如mean為theta new的正常分配)

<https://github.com/EnChiSu/Master-Thesis/blob/master/Check%20Point%202/Evaluating%20MCMC%20performance%20%26%20model%20fit.md>

延伸：為什麼會說窄的proposal的acceptance rate較高，從圖形中不是寬的proposal也涵蓋滿大的面積嗎？



- MCMC的acceptance rate跟rejection method的acceptance rate在概念上是有差異的。
- MCMC說的acceptance rate是你隨著Markov Chain每一期移動，整體模擬過程下每一步被接受的比率，是一個動態的過程。寬的proposal之所以接受率低是因為在MCMC的過程他很快就跳到目標分配的中心 (每一個步伐都很大)，在這之後就不太跳了，一直定在這個位置，造成後面每一次都reject，所以整體接受率低。
- Rejection method的acceptance rate就如前面投影片的图片，受到目標分配在envelope中所含蓋的面積比率決定，為一個靜止狀態的衡量指標。

問題二：  
如何用MCMC求出的posterior做prediction

## 問題二：來自Posterior predictive distribution的想法

可以想像為 likelihood, 為 base on posterior distribution 的 likelihood

$$p(y^{rep} | y) = \int p(y^{rep} | \theta) p(y | \theta) p(\theta) d\theta$$

(y rep denote future observation)

$$p(\theta | \text{data}) \propto p(\text{data} | \theta) p(\theta).$$

可以想像為 prior x likelihood 中的 prior, 只不過這邊使用經計算出來的 posterior

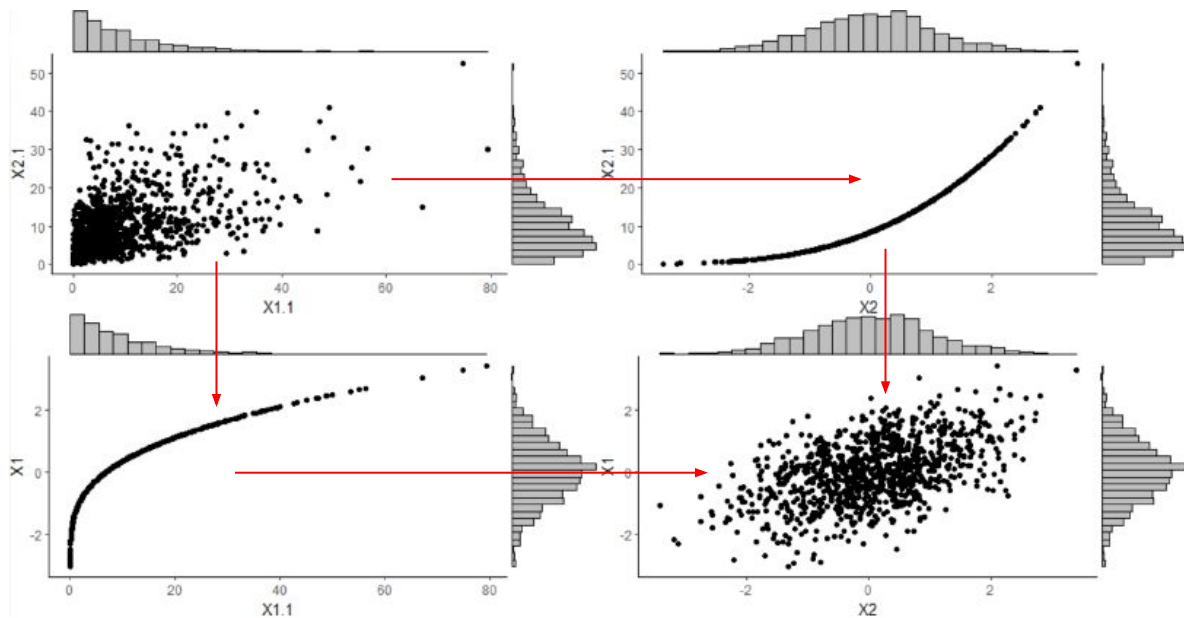
我們可以透過 MCMC 找到 posterior 參數。那我們要預測是要用這個 posterior 的 mean 或是如果是 gamma 就用它的 rate 參數作為預測嗎？

另外如果將觀察到的資料帶入這個求出來的 posterior 當中得到的機率是 likelihood 不依定會是最大因為我們是用 MCMC 而不是 MLE 做估計，這樣理解是對的嗎？這個帶資料求出來的 likelihood 可以協助我們預測嗎？

問題三：  
如何使用Gibbs Sampling估correlation "matrix"



# 問題三：使用Gibbs Sampling估correlation matrix



Note: 此連結為透過 gaussian copula 轉 bivariate normal 到 bivariate gamma 的模擬示範  
<https://github.com/EnChiSu/Master-Thesis/blob/master/Check%20Point%202/Gaussian%20Copula.ipynb>

- 來由：  
在學長的論文當中，因為不同商品的購買間格時間服從 multivariate gamma distribution，且變數間存在相關性(實證資料如左上圖)，但因為為 gamma 不好求當中的相關性。故透過 gaussian copula 的方式轉成 multivariate standard normal，在用 gibbs sampling 的方式求得 correlation matrix
- 問題：  
bivariate 下 correlation 為一個值，若為 multivariate 下會變為 correlation matrix，Gibbs Sampler 如何去估計？還是需要針對每個倆倆變數去估計相關係數值？