UiO : Matematisk institutt

Det matematisk-naturvitenskapelige fakultet

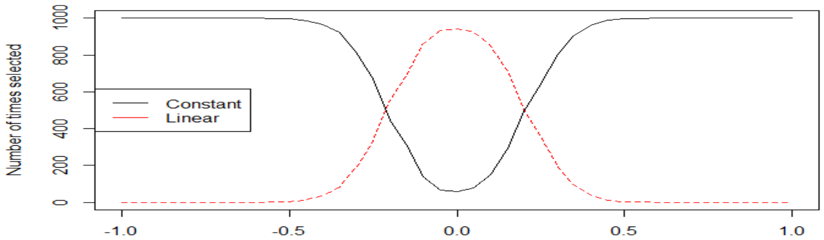
STK-4051/9051 Computational Statistics Spring 2020
Comments to exercise 9

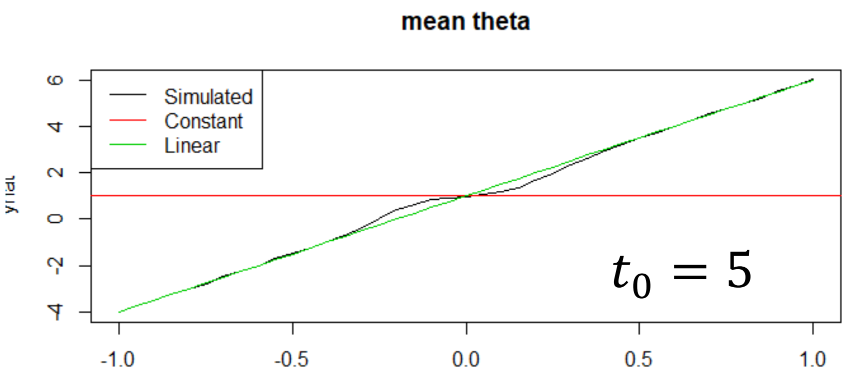
Instructor: Odd Kolbjørnsen, oddkol@math.uio.no

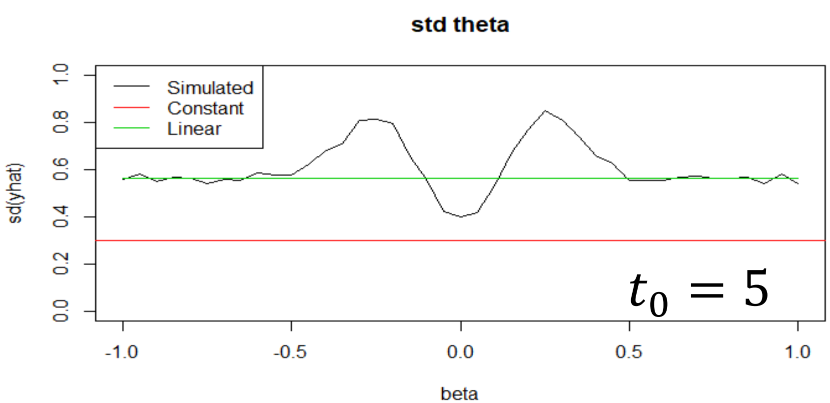
Ex 32

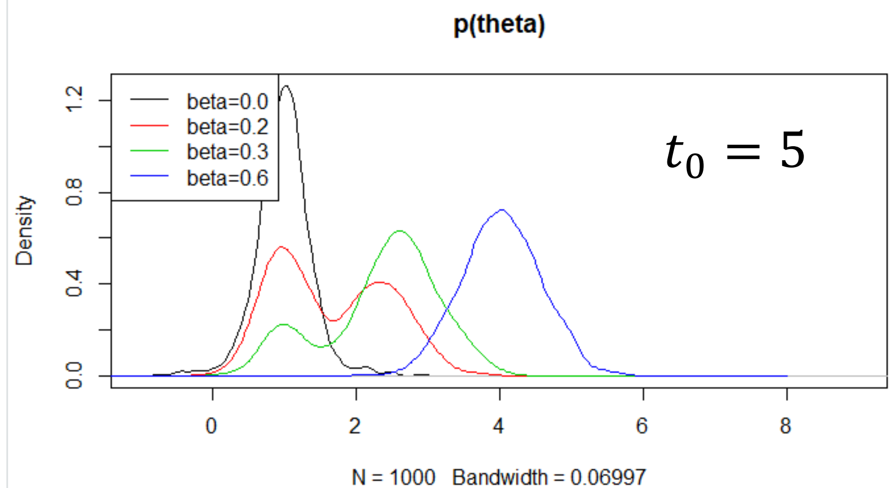
This task is just an example of how you can use simulations to check the method you are using. In this case for model selection.

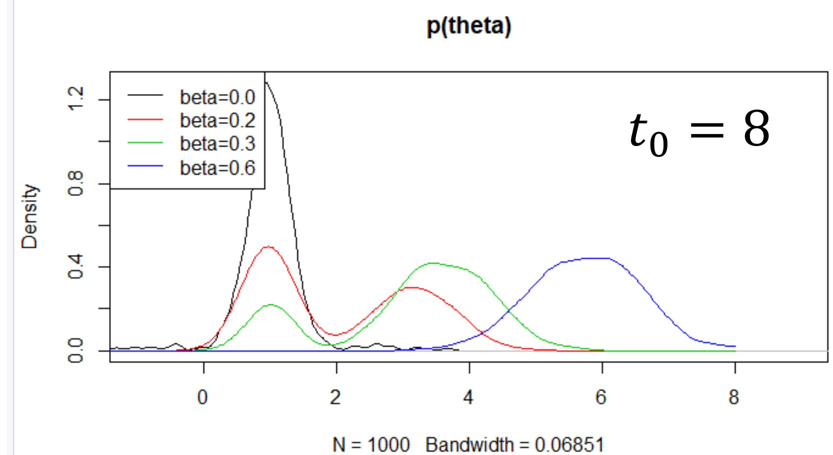
a) See code in R-file for implementation





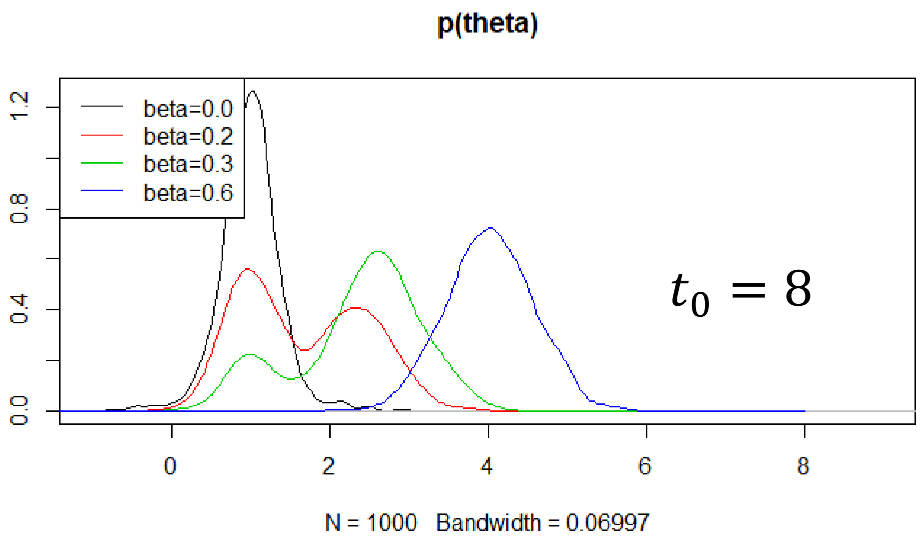




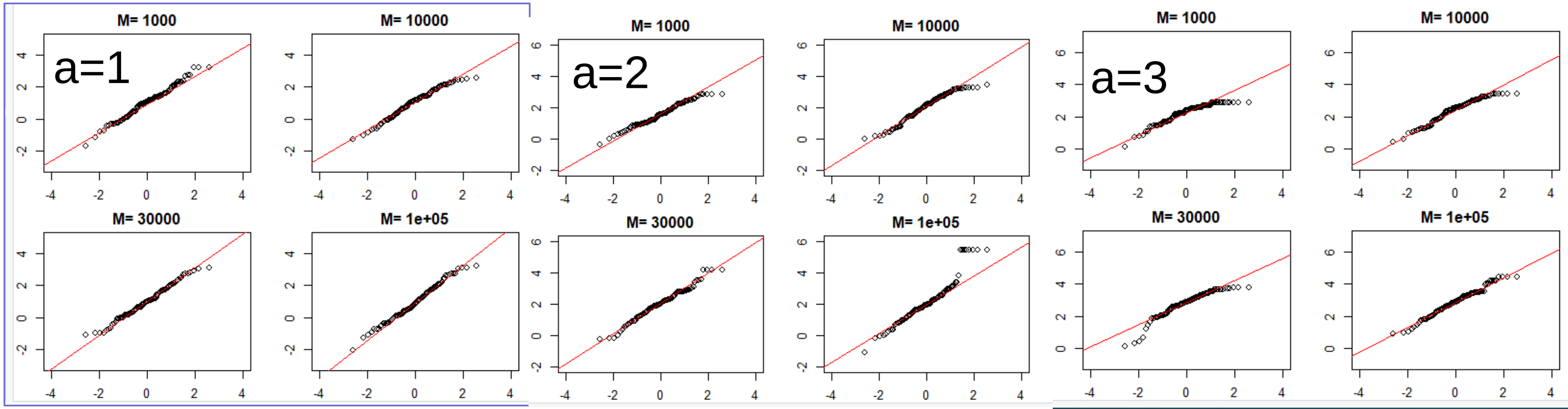


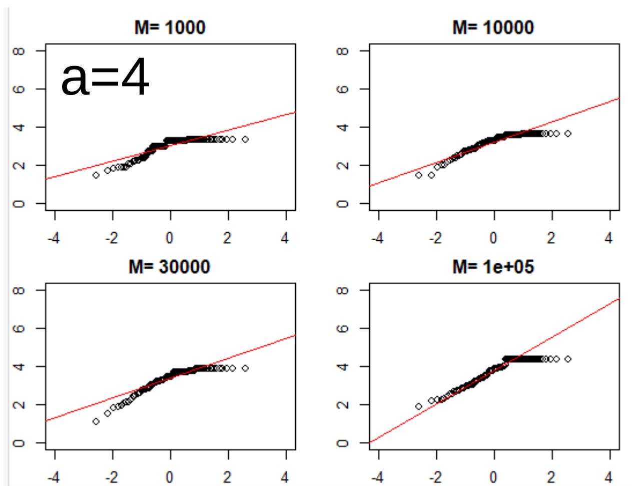
Comments

- The standard approach does not propperly take the uncertainty into account when it is used for model selection.
- The true model is acctually bi-modal.



Ex 33 importance sampling





c) As a increase (1 to 4):

- The effective number of samples goes down
- The qq plot does not match the line
- The mean is too low
- The std is too low

d)

- The weigths are normalized prior to re-sampling

Ex 35

- Random Walk
- Use library mvtnorm
- B) We need more than 1000 samples to get good estimates, but we are in the ball park
- C) We are not even close with the 1000 samples.
  We are trapped in the valley.

$$R(\mathbf{x},\mathbf{x}^*) = \frac{f(\mathbf{x}^*)g(\mathbf{x}|\mathbf{x}^*)}{f(\mathbf{x})g(\mathbf{x}^*|\mathbf{x})} = \frac{f(\mathbf{x}^*)}{f(\mathbf{x})}$$

```
acc = 0
for(i in 2:N)
{
  j = sample(1:2,1)
  y = X[i-1,]
  y[j] = rnorm(1,X[i-1,j],sig.prop)
  R = dmvnorm(y,mu,Sigma)/dmvnorm(X[i-1,],mu,Sigma)
  if(runif(1)<R)
  {
    X[i,] = y
    acc = acc +1
  }
  else
    X[i,] = X[i-1,]
}
```

Ex 36

Exercise 36 (Sampling random variables by Metropolis)
Let X be a random variable with density $f(x) = ch(x)$, where h is a computable expression. We shall in this exercise discuss sampling of X through the Metropolis algorithm. The aim is to demonstrate how the algorithm works, not to produce the most efficient sampling scheme conceivable.

Let $\varepsilon_1, \varepsilon_2, ..$ be a sequence of independently and identically distributed random variables. You may choose any distribution symmetrical about the origin. Suggested choice: The uniform over $(-a, a)$, where $a > 0$ has to be adapted for each application. In addition, let $U_1, U_2, ...$ be a sequence of uniforms over $(0, 1)$. Consider the recursion

$$X_{t+1} = X_t + \varepsilon_t I_t$$

where

$$I_t = 1, \qquad \text{if } U_t \leq \frac{h(X_t + \varepsilon_t)}{h(X_n)}$$

and $= 0$ otherwise.

(a). Show that this recursion is a special case of the Metropolis algorithm. Explain, in particular, the relevance of demanding ε_t to be a symmetric distribution.

Solution to exercise 36. (a). We have that the proposal is $X_t^* = X_t + \varepsilon_t$. Since the proposal distribution is symmetric, the Metropolis-Hastings ratio becomes

$$R_t = \frac{h(X_t^*)}{h(X_t)} = \frac{h(X_t + \varepsilon_t)}{h(X_t)}.$$

If we generate $U_t \sim \text{Uniform}[0, 1]$, then we can accept if $U_t \leq \min\{1, R_t\}$ which is equivalent to accept if $U_t < R_t$. This means that

$$X_{t+1} = \begin{cases} X_t^* & \text{if } U_t < R_t \\ X_t & \text{otherwise} \end{cases}$$
$$= X_t + I_t \varepsilon_t$$

given the definition of I_t.

Making the distribution of ε_t to symmetric simplifies the MH-ratio in that the proposal densities disappear.

(b). Suppose f is the standard normal density. Show that the ratio defining I_t becomes $\exp(-X_t\varepsilon_t - 0.5\varepsilon_t^2)$. I_t is more likely to become 1 if the signs of X_t and ε_t are alternate. Explain the meaning of this.

(c). Will you need all U_t to run the algorithm?

(b). In this case

$$R_t = \frac{\exp(-0.5(X_t + \varepsilon_t)^2)}{\exp(-0.5X_t^2)} = \exp(-X_t\varepsilon_t - 0.5\varepsilon_t^2)$$

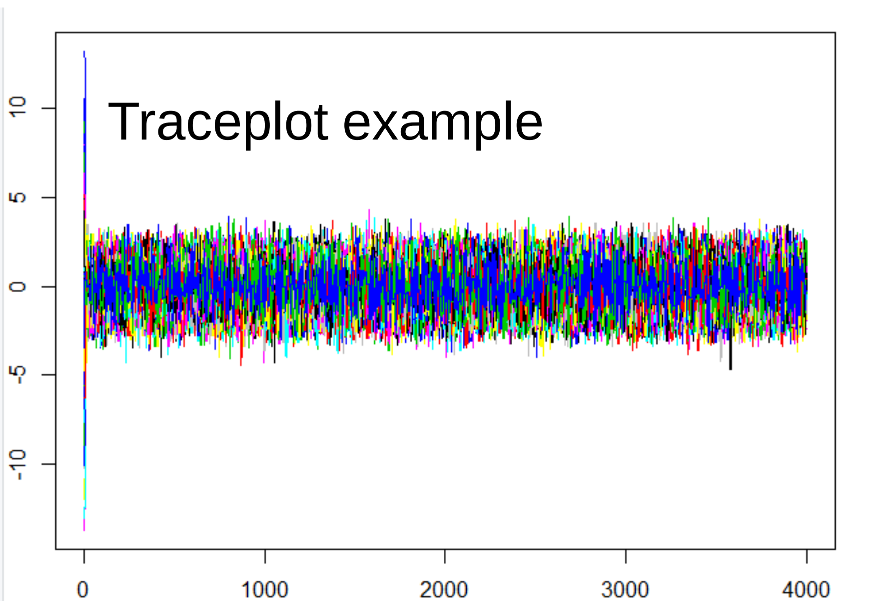
Since we have a distribution centered at zero, we would like to mostly move towards zero, which means making ε_t negative if X_t is positive and vice verse.

(c). No, if $R_t \geq 1$, you do not need to generate U_t.

Ex 36

d)

```
a = 5
acc = 0
for(i in 2:N)
{
  eps = runif(m,-a,a)
  R = dnorm(X[i-1,]+eps)/dnorm(X[i-1,])
  I = as.numeric(runif(m)<R)
  X[i,] = X[i-1,]+I*eps
  acc = acc + sum(I)
}
```



- f) use multiple start points, (extreme values or random)
- e,g,h) This is by trial and error, do it
  I will discuss similar issues in the lecture probably in two weeks
- i) Changing *a* you do not have detailed balance, since the reverse probability is different This must be accounted for