MH4500 LAB 2

Kan Kar Shen
U1840283A

Q1

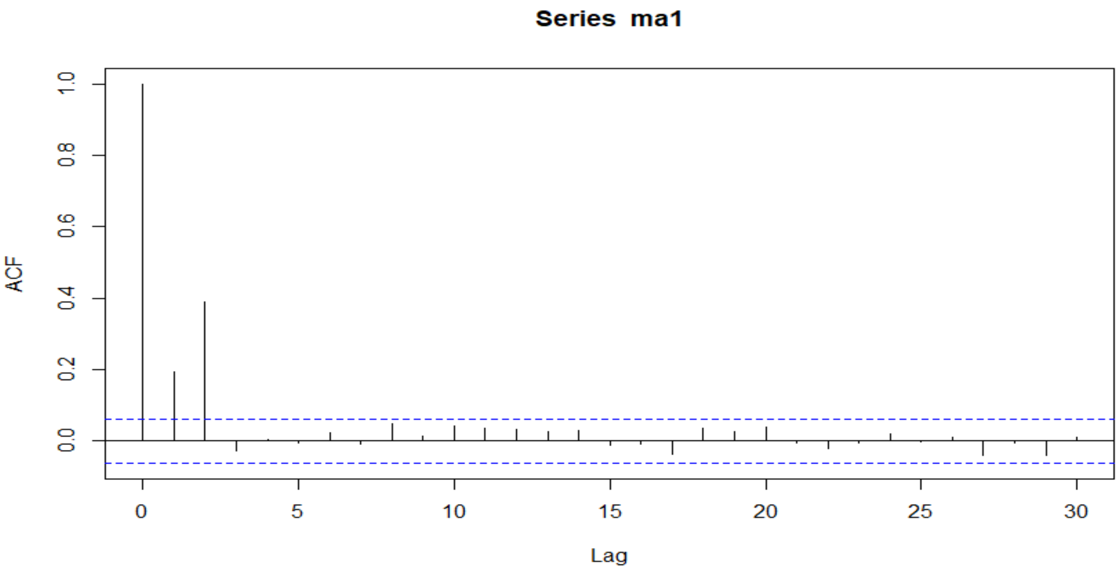
```
masim = function(theta, sigsq, T) {
  q = length(theta) #number of lags in AR
  noise = rnorm(T+q, sd = sqrt(sigsq)) #generate white noise and afew to start
  x = c(noise[1:q],rep(0,T)) #put initial noise term in, setting rest to 0
  for( i in (q+1):(T+q)) { #generate AR
  x[i] = theta %*% noise[i - (1:q)] + noise[i]
  }
  x = x[(q+1):(T+q)] # remove initial starting positions
  x #return time series
}
```

```
> masim = function(theta, sigsq, T) {
+    q = length(theta) #number of lags in AR
+    noise = rnorm(T+q, sd = sqrt(sigsq)) #generate white noise and afew to start
+    x = c(noise[1:q],rep(0,T)) #put initial noise term in, setting rest to 0
+    for( i in (q+1):(T+q)) { #generate AR
+    x[i] = theta %*% noise[i - (1:q)] + noise[i]
+    }
+    x = x[(q+1):(T+q)] # remove initial starting positions
+    x #return time series
+ }
> 
```

Q2

T = 1000,

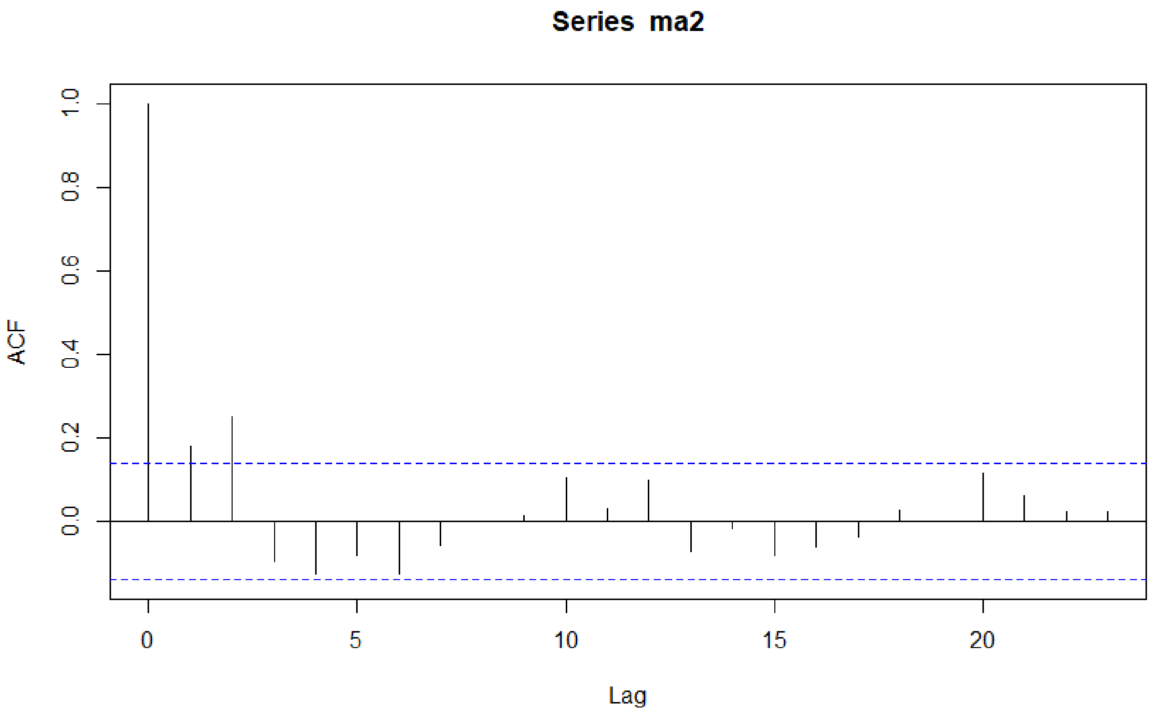
```
> ma1 = masim(c(0.5,2),1,1000)
> acf(ma1)
> 
```



Yes, the generated result is consistent with the model generated as the ACF of a MA(q) time series cut off after lag q. The generated model cuts off at lag 2 which is consistent with the MA(2) model. In addition, there is a positive correlation until lag 2 as its above the blue dotted line, implying significance.

When T = 200,

```
> ma2 = masim(c(0.5,2),1,200)
> acf(ma2)
>
```



The blue dotted lines shows the point of significance, where falling beyond the blue dotted lines, the autocorrelations is implied to be statistically significantly different from zero
The range of the blue dotted lines are wider for the ma2 model, which has a smaller sample size, as compared to the ma1 model, which has a larger sample size.

The autocorrelations change for every iteration and they always cut off at lag 2.

Q3

There are 2 datasets with different number of observations, one with T = 1000 and T = 200 observations.

For T = 1000,

```
install.packages("forecast")
library(forecast)
fit_ma1 = auto.arima(ma1)
summary(fit_ma1)
```

```
> summary(fit_ma1)
Series: ma1
ARIMA(0,0,3) with zero mean

Coefficients:
         ma1     ma2     ma3
      0.3092  0.4951  0.0503
s.e.  0.0311  0.0284  0.0323

sigma^2 estimated as 3.851:  log likelihood=-2091.83
AIC=4191.67   AICc=4191.71   BIC=4211.3

Training set error measures:
                     ME     RMSE      MAE      MPE     MAPE      MASE         ACF1
Training set 0.03753317 1.959346 1.561144 44.68776 160.2873 0.7561869 0.0007197302
```

After fitting the ARIMA model for the T = 1000 dataset, the best fitting model is ARIMA(0,0,3) with zero mean

For T = 200,

```
> fit_ma2 = auto.arima(ma2)
> summary(fit_ma2)
Series: ma2
ARIMA(2,0,2) with zero mean

Coefficients:
          ar1     ar2     ma1     ma2
      -0.0779  0.0544  0.2928  0.5615
s.e.   0.1472  0.1743  0.1335  0.1504

sigma^2 estimated as 3.949:  log likelihood=-419.56
AIC=849.12   AICc=849.43   BIC=865.61

Training set error measures:
                     ME     RMSE      MAE      MPE     MAPE    MASE        ACF1
Training set 0.02332689 1.967291 1.579937 421.5516 456.5141 0.69709 0.008345777
```

After fitting the ARIMA model for the T = 200 dataset, the best fitting model is ARIMA(2,0,2) with zero mean