Time-Series-Assessment-2.R

Grace Laryea

2022-04-03

```
#Code is in deep blue

#Loading of library packages to enable the functionality of my functions
library(tseries)

## Warning: package 'tseries' was built under R version 4.0.5

## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo

library(forecast)

## Warning: package 'forecast' was built under R version 4.0.5

Q1
#..........................................................................#

#(a)
#Setting random number generator at value 639
set.seed(639)
#Running ARIMA Simulation
#n= number of values, 350 values in this case
#model list defines parameters of AR and MA series
#alpha 1=0.4, alpha 2=-0.1 for AR series
#beta1 =0.57
x=arima.sim(n=350,model=list(ar=c(0.4,-0.1),ma=0.57))
```

....

Explanation:

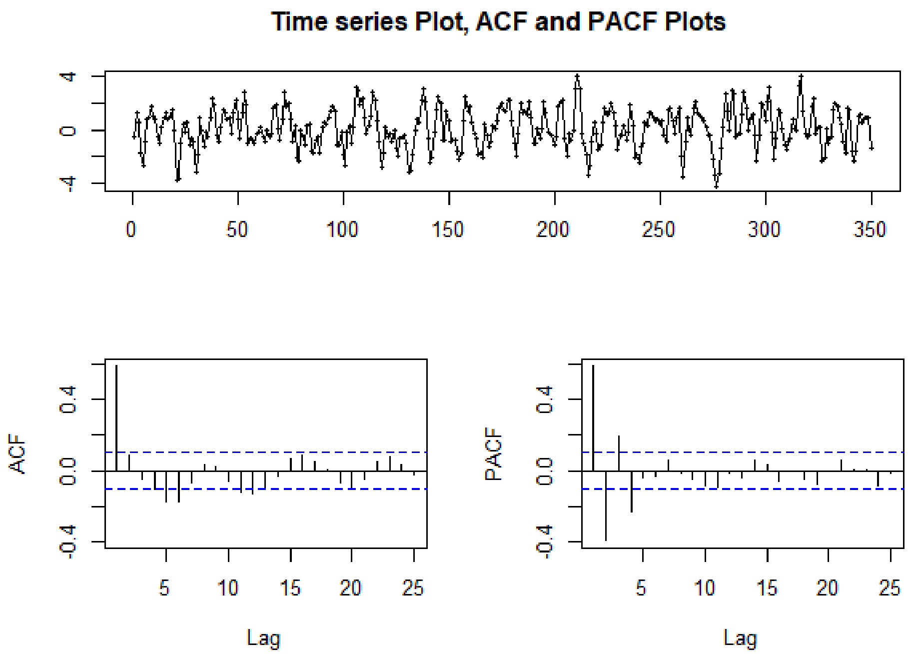
We are simulating a time series of length 350, and ARMA(2,1) with parameters $\alpha_1 = 0.4, \alpha_2 = -0.1\ and\ \beta_1 = 0.57\ respectively$

Also, the set.seed, ensure we are able to replicate our time series each time , without randomly generating different values each time.

....

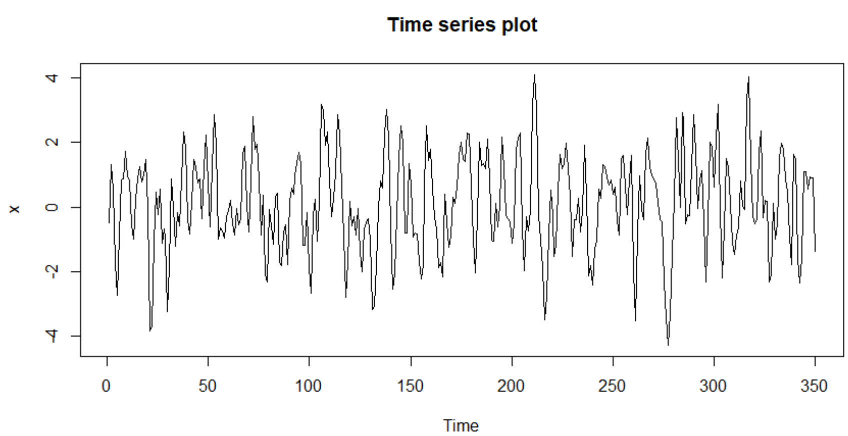
```r
Q1(b)
#We use the tsdisplay function to plot the ARIMA time series simulation of x
#This produces a graph of the Time series, its respective ACF and PACF graphs

#code
tsdisplay(x,main="Time series Plot, ACF and PACF Plots ")
```



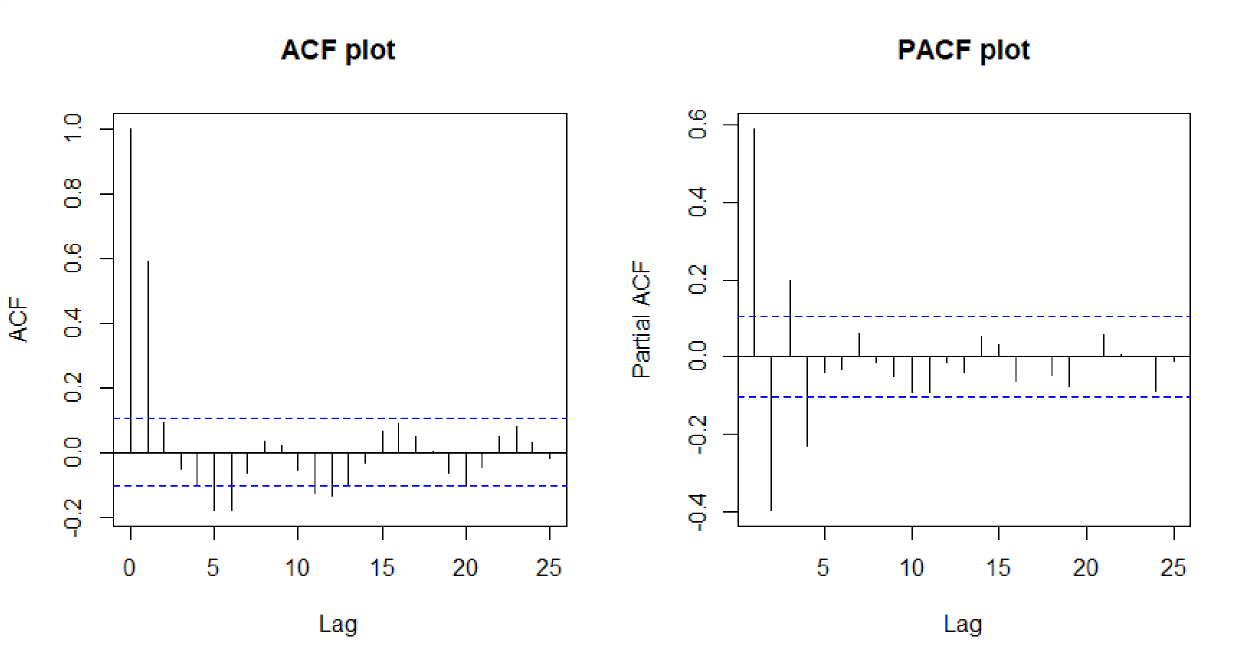

```r
Q1b)a #Can also be produced using ts.plot(), acf() and pacf()
ts.plot(x)
```

Output



Q1b)b

```r
#code:
par(mfrow=c(1,2))
acf(x)
pacf(x)
```



```
Q1(c)
Our simulation of Time series is stationary, but I am checking for clarity
Running Dicky-Fuller Test to prove stationarity

#code:
adf.test(x)
```

Output

```
## Warning in adf.test(x): p-value smaller than printed p-value

##
##  Augmented Dickey-Fuller Test
##
## data:  x
## Dickey-Fuller = -6.695, Lag order = 7, p-value = 0.01
## alternative hypothesis: stationary

Q2c)a
#Fitting the ARIMA time series to get ARMA(2,2)
#ARIMA(2,0,2) or ARMA(2,2)
#I have assigned ARMA(2,2) as fit 1
fit1=arima(x,order=c(2,0,2))
#Calling fit1
```

Output

```
fit1

##
## Call:
## arima(x = x, order = c(2, 0, 2))
##
## Coefficients:
##           ar1      ar2     ma1     ma2  intercept
##       -0.0655  -0.0207  1.0440  0.3108     0.0982
## s.e.   0.6073   0.2067  0.6063  0.4021     0.1229
##
## sigma^2 estimated as 1.129:  log likelihood = -518.38,  aic = 1048.77

#(b)
#ARIMA(2,0,1) or ARMA(2,1)
#Assigning ARMA(2,1) as fit2
fit2=arima(x,order = c(2,0,1))
#Calling fit2
```

Output

```
fit2

##
## Call:
## arima(x = x, order = c(2, 0, 1))
##
```

```
## Coefficients:
##          ar1      ar2     ma1  intercept
##       0.3775  -0.1402  0.5969     0.0986
## s.e.  0.0832   0.0726  0.0694     0.1188
##
## sigma^2 estimated as 1.131:  log likelihood = -518.64,  aic = 1047.28
```

Q1(c)

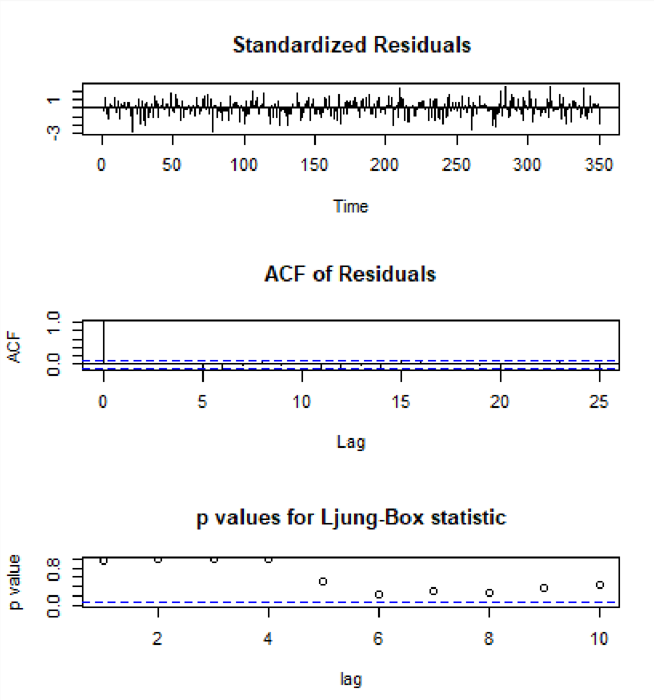
I choose model 2, ARMA(2,1) as it has the lowest AIC, Maximum likelihood value. Hence AR MA(2,1) is the best model.

ARMA(2,2)--→ log likelihood = -518.38, aic = 1048.77

ARMA(2,1)--→log likelihood = -518.64, aic = 1047.28

```
Q1(d)
#Running a diagnostic test of fit2
tsdiag(fit2)
```

Output



Equation of model chosen:

ARMA(2,1)

$y_t = 0.4y_1 - 0.1y_2 - 0.57\varepsilon_1$

The output suggests some forecasting can be done, also, the ACF plots suggest white noise a s well as the p-values of Ljung-Box statistic being greater than the significance value of 5%.

Also, since our p-values are greater than 5%, we can suggest, our null hypothesis is not true , hence no evidence of serial correlation amongst the fitted model (fit2).

Looking at the 2nd plot, we can observe the lags of the ACF residuals do not exceed the 95% level, hence model is quite a good fit!

We will look into this further in part(e)

```
Q1(e)
# I assumed my x value to be my chosen fitted ARIMA series
#fitdf= number of degrees of freedom to be subtracted if x is a series of
#residuals.
#In our case its 4, as our p=2 from AR terms and q=1 from MA terms and -1
#We normally don't tend to use it, but as I am using my fitted time series to
#check for serial correlation
#The question demanded a lag of length 10
#Type of Box.test is Ljung-Box
Box.test(residuals(fit2),fitdf =4,lag = 10,type = "Ljung-Box")
```

Output

```
##
##  Box-Ljung test
##
## data:  residuals(fit2)
## X-squared = 9.8832, df = 6, p-value = 0.1297
```

....

Explanation:

$$Q = T(T+2)\sum_{k=1}^{s} r_k^2/(T-k) \rightarrow \chi^2_{s-p-q} \qquad (4.11)$$

X-squared represents the Q value for Ljung-Box test(Portmanteau statistic),so our Q value is 9.8832, total degree of freedom= 10-3-1=6,p-value= is probability of our Q value occurring.

So in this context:

Our p-value > 5%, hence adequacy of fitted ARMA(p,q) should be re-analyzed.

$H_0 = Residuals\ of\ white\ noise\ observed$

$H_1 = No\ residuals\ of\ white\ noise\ observed$

T= 350

```
#using the checkresiduals() to see if my answer is right
checkresiduals(fit2)
```

Output

```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(2,0,1) with non-zero mean
```

```
## Q* = 9.8832, df = 6, p-value = 0.1297
##
## Model df: 4.   Total lags used: 10

#Since p-value is greater than 5%, we need to use the GARCH model

#.........................................................................#
```

Question 2

Generate a random sample of y!

a) $g(y) = \frac{1}{2}\sin(y) \quad ; 0 < y < \pi$

$$G_Y(u) = \int_{-\infty}^{\infty} g_Y(u)\, dy$$

$$= \int_0^{u} \tfrac{1}{2}\sin(y)\, dy$$

$$= \tfrac{1}{2}\int_0^{u} \sin(y)\, dy$$

$$= \tfrac{1}{2}\Big[-\cos(y)\Big]_0^{u}$$

$$= \tfrac{1}{2}\Big[\cos(y)\Big]_u^{0}$$

$$= \tfrac{1}{2}\Big[1 - \cos(u)\Big]$$

$$G_Y(u) = \tfrac{1}{2}(1 - \cos(u))$$

Making u the subject!

$$G_Y(y) = y$$

$$y = \tfrac{1}{2}(1 - \cos(u))$$

$$2y = 1 - \cos(u)$$

$$1 - 2y = -\cos(u)$$

$$2y - 1 = \cos(u)$$

$$\arccos(2y-1) = u$$

$$\therefore u = \arccos(2y-1)$$

b) Method of rejection sampling

$$Y \sim g(y) = \tfrac{1}{2}\sin(y)$$

$$X \sim f(x) = \frac{1}{\pi} \quad \} \; X \sim U(0,\pi)$$

We know the $G_Y^{-1}(u) = \arccos(2u-1)$

$X = Y$

$$\sup_y \frac{g(y)}{f(y)} = M$$

$$M = \sup_y \frac{g(y)}{f(y)}$$

$$= \frac{\tfrac{1}{2}\sin(y)}{\tfrac{1}{\pi}}$$

$$= \tfrac{\pi}{2}\sin(y) \Rightarrow \ln\left(\tfrac{\pi}{2}\sin(y)\right)$$

$$\frac{d}{dy}\left[\ln\left(\tfrac{\pi}{2}\sin(y)\right)\right] = \frac{\cos(y)}{\sin(y)}$$

$$\frac{\cos(y)}{\sin(y)} = 0$$

$$\cos(y) = 0$$

$$y = \arccos(0)$$

$$y = \pi/2 = n\pi/2 \quad \text{for } n = 1, 2, 3, \ldots$$

$$\frac{d^2y}{dy^2} = \frac{-\cos^2(y)}{\sin^2(y)} - 1 < 0$$

$\frac{d^2y}{dy^2} < 0$, hence $y = \pi/2$ is a maximum turning point.

$$M = \sup_y \left(\frac{g(y)}{f(y)}\right)$$

$$= \left(\tfrac{1}{2} \sin(\pi/2) * \pi\right)$$

$$= \pi/2 \sin(\pi/2)$$

$$\therefore \frac{g(y)}{Mf(y)} = \frac{1/2 \sin(y)}{\pi/2 \sin(\pi/2)\, 1/\pi} = \frac{\frac{\sin(y)}{2}}{\frac{1}{\pi} * \pi/2 \sin \pi/2}$$

$$= \frac{\sin(y)/2}{\frac{1}{2} \sin \pi/2} = \frac{\sin(y)}{\cancel{2}} * \cancel{2} * \frac{1}{\sin(\pi/2)}$$

$$\frac{g(y)}{Mf(y)} = \frac{\sin(y)}{\sin(\pi/2)} = \sin(y)$$

Since $\sin(\pi/2) = 1$

Algorithm

1. Generate a random number U_1
2. Set $Y = \arccos(2u-1)$ [Solved in (a) by inversion]
3. Generate another random number U_2
4. If $U_2 \leq \sin(y)$, then set $X = Y$,

   or otherwise start all over from 1

Question 3

$$f(x) = \frac{e^{-x^2/2}}{(\Phi(3) - 0.5)\sqrt{2\pi}} \quad , \quad 0 < x < 3$$

$$g(y) = 2\alpha e^{-2y} \quad , \quad 0 < y < 3$$

(a) $$\int_{-\infty}^{\infty} g_{y}(y)\, dy = 1$$

$$\int_0^3 2\alpha\, e^{-2y}\, dy = 1$$

$$\left[\frac{2\alpha e^{-2y}}{-2} \right]\Bigg|_0^3 = 1$$

$$\left[-\alpha e^{-2y} \right]\Bigg|_0^3 = 1$$

$$\left[\alpha e^{-2y} \right]\Bigg|_3^0 = 1$$

$$\alpha e^{-0} - \alpha e^{-6} = 1$$

$$\alpha - \alpha e^{-6} = 1$$

$$\alpha(1 - e^{-6}) = 1$$

$$\alpha = \frac{1}{1 - e^{-6}}$$

$$\therefore \alpha = (1 - e^{-6})^{-1}$$

Hence distribution function for Y is Exp (2)

Since $(1-e^{-6})^{-1} \approx 1$

$$p.d.f = \begin{cases} 2(1-e^{-6})^{-1}e^{-2y} & , 0 < y < 3 \\ 0 & , \text{otherwise} \end{cases}$$

b)

Algorithm for generating samples of random variable Y using standard uniform $U(0,1)$

$Y \sim Exp(2)$

$$G_Y(u) = \int_{-\infty}^{u} g_Y(y)\, dy$$

$$= \int_0^u g_Y(u)\, dy$$

$$= \int_0^u 2(1-e^{-6})^{-1} e^{-2y}\, dy$$

using integral by substitution:

$$\frac{2}{1-e^{-6}} \int_0^u e^{-2y}\, dy$$

let $u = -2y$

$$\frac{du}{dy} = -2$$

$$-\frac{du}{2} = dy$$

$$\frac{\cancel{2}}{1-e^{-6}} \int_0^u e^u \frac{du}{-\cancel{2}} \Rightarrow \frac{-1}{1-e^{-6}} \int_0^u e^u\, du$$

$$\frac{-1}{1-e^{-6}} \int_0^u e^u \, du$$

$$\frac{-1}{1-e^{-6}} \left[e^u \right]\Big|_0^u$$

$$\frac{-1}{1-e^{-6}} \cdot (e^u - e^0) = \frac{-e^u + e^0}{1-e^{-6}} = \frac{1-e^u}{1-e^{-6}}$$

$$= \frac{1-e^{-2y}}{1-e^{-6}} \quad \text{since} \quad u = -2y$$

$$\therefore \quad G_Y(u) = \frac{1-e^{-2y}}{1-e^{-6}} \qquad \text{let } G_Y(y) = u$$

Setting: $\dfrac{1-e^{-2y}}{1-e^{-6}} = u$

$$1 - e^{-2y} = u(1-e^{-6})$$

$$1 - u(1-e^{-6}) = e^{-2y}$$

$$\ln\left[1 - u(1-e^{-6})\right] = -2y$$

$$y = -\frac{\ln\left[1-u(1-e^{-6})\right]}{2}$$

$$y = -\frac{\ln\left[1-u(1-e^{-6})\right]}{2}$$

(c)

X has p.d.f of $f(x)$, and the inversion is hard.
Y has p.d.f of $g(x)$, and the inversion is easy.

Lets assume;

$$\frac{f(x)}{g(x)} \leq M < \infty \quad \text{for all values of } x.$$

Generate a random variable Y with p.d.f $g(y)$.

$$f(x) = \frac{e^{-x^2/2}}{(\Phi(3) - 0.5)\sqrt{2\pi}}$$

$$g(x) = 2(1-e^{-6})^{-1}e^{-2x}$$

let $x = Y$

$$M = \sup_x \frac{f(x)}{g(x)}$$

We will use MLE. to get a proper estimate of the maximum value of x;

$$= \frac{e^{-x^2/2}}{(\Phi(3) - 0.5)\sqrt{2\pi}} * \frac{(1-e^{-6})}{2e^{-2x}}$$

$$= \frac{e^{-x^2/2 + 2x} * (1-e^{-6})}{2(\Phi(3) - 0.5)\sqrt{2\pi}}$$

$$\ln\left[\frac{e^{-x^2/2+2x}(1-e^{-6})}{(\Phi(3)-0.5)\sqrt{2\pi}(2)}\right]$$

$$-\frac{x^2}{2}+2x-\ln\left[(\Phi(3)-0.5)\,2\sqrt{2\pi}\right]+\ln\left[1-e^{-6}\right]$$

$$\frac{d}{dx}\left[\frac{-x^2}{2}+2x\right]-\frac{d}{dx}\left[(\Phi(3)-0.5)\,2\sqrt{2\pi}\right]+\frac{d}{dx}(1-e^{-6})$$

$$\frac{d}{dx}=\frac{-2x}{2}+2$$

set to 0.

$$\frac{d}{dx}\stackrel{!}{=}0, \quad \frac{-\cancel{2}}{\cancel{2}}x+2=0, \quad -x+2=0$$

$$\therefore \quad \underline{\underline{x=2}} \quad \text{(Turning point)}$$

Maximum of x occurs at 2

$$\frac{d^2}{dx^2}=-1 \quad \text{(Hence maximum occurs at } x=2\text{)}$$

$$\frac{f(x)}{Mg(x)}=\frac{e^{-x^2/2+2}(1-e^{-6})}{2(\Phi(3)-0.5)\sqrt{2\pi}\cdot 2}$$

$$=\underline{\underline{\frac{e^{-x^2/2+2+2x}(1-e^{-6})}{4\sqrt{2\pi}(\Phi(3)-0.5)}}}=e^{-x^2/2+2x+2}$$

$$M = \frac{e^{-x^2/2 + 2x}(1-e^{-6})}{2(\Phi(3)-0.5)\sqrt{2\pi}}, \text{ where } x=2$$

$$M = \frac{e^{-2^2/2 + 2(2)}(1-e^{-6})}{2(\Phi(3)-0.5)(\sqrt{2\pi})}$$

$$M = \frac{e^{-2+4}(1-e^{-6})}{2(\Phi(3)-0.5)(\sqrt{2\pi})}$$

$$= \frac{e^{+2}(1-e^{-6})}{2(\Phi(3)-0.5)\sqrt{2\pi}}$$

$$\frac{f(x)}{Mg(x)} = \frac{e^{-x^2/2}\cancel{(1-e^{-6})}e^{2x}}{\cancel{2\sqrt{2\pi}(\Phi(3)-0.5)}} \times \frac{\cancel{2\sqrt{2\pi}(\Phi(3)-0.5)}}{e^2\cancel{(1-e^{-6})}}$$

$$= e^{-x^2/2+2x-2}$$

$$= e^{-x^2/2+2x-2}$$

Algorithm:

① Generate a random number U_1

② Set $Y = -\frac{1}{2}\ln[1-u(1-e^{-6})]$

③ Generate another random number U_2

④ If $U_2 \leq e^{-y^2/2+2y-2}$, then set $X=Y$. Otherwise Go to step ①.

c) Algorithm to perform run in (c)

① Generate a random number U_1

$U_1 = 0.86$

② Set $Y = -\frac{1}{2}\ln[1 - u_1(1-e^{-6})]$:

$Y = -\frac{1}{2}\ln[1 - 0.86(1-e^{-6})]$

$Y = 0.9755$

③ Generate another random number U_2

$U_2 = 0.34$, where $h(y) = e^{-y^2/2 + 2y - 2}$

④ If $U_2 \leq e^{-y^2/2 + 2y - 2}$

$U_2 \leq e^{-0.9755^2/2 + 2(0.9755) - 2}$

$U_2 \leq 0.59167$

Since $U_2 \leq 0.59167$, we ~~Go to step 1~~ end here!

~~and start the process over!~~

So $u \leq h(y) \Rightarrow Mg(y)u \leq f(y)$

where $Mg(y)u$ is a random point below $Mg(y)$),

~~Hence we accept this point below f(y)).~~

(d) Z = number of attempts until we accept X

p = Probability (Accept X at each attempt)

$$p = P(\text{Accept } X \text{ at each attempt}) = \int_{-\infty}^{\infty} h(y)\, g(y)\, dy$$

$$= \int_{-\infty}^{\infty} \frac{f(y)}{Mg(y)}\, g(y)\, dy$$

$$= \frac{1}{M} \int_{-\infty}^{\infty} f(y)\, dy$$

$$= \frac{1}{M}$$

$$M = \frac{e^{2}(1-e^{-6})}{2(\sqrt{2\pi})(\Phi(3)-0.5)}$$

$$\frac{2\sqrt{2\pi}\ (\Phi(3-0.5))}{e^{2}(1-e^{-6})}$$

$$p = \frac{2\sqrt{2\pi}\ (\Phi(3)-0.5)}{e^{2}-e^{-4}}$$

//

$$\Phi(3) = 0.9986501$$

$$p = 0.3392$$

$\mathbb{E}(z)$ = number of pairs of random numbers.

$= \frac{1}{p} = M$

$= 2.948$

$= 2.95 \approx 3$ pairs of Random numbers.