Roll No-12

M.sc. 3rd semester

Date of Assignment-08/12/2020

Date of Submission-14/12/2020

Experiment No -10

Topic- Calculation of Auto covariance and Auto correlation function for time series analysis

Problem – Following are the temperature measurements made on a chemical reactor every minute. Calculate the Auto covariance and Auto correlation function up to lag 5.

Temperature-
200,202,208,204,202,201,200,199,201,198,200,203,201,211,204,206,203,203,204,207,206,207, 206,207,206,200,207,205,200,195,202, 204,203,198,200.

Theory and Calculation-

The auto-covariance at lag k is given by-

$$R_k = \text{cov}(Z_t, Z_{t-1})$$

$$=> R_k = E\left[\{Z_t - E(Z_t)\}\{Z_{t+1} - E(Z_{t+1})\}\right] \text{ -----(1)}$$

We know for a stationary time series -

$$E(Z_t) = E(Z_{t+1}) = \mu \quad ; \quad k = 1,2,3,.......... \text{ ------(a)}$$

$$V(Z_t) = V(Z_{t+1}) = \sigma^2 \quad ; \quad k = 1,2,3,.......... \text{ ------(b)}$$

Therefore, equation (1) becomes-

$$=> R_k = \text{cov}(Z_t, Z_{t-1}) = E\left[\{Z_t - \mu\}\{Z_{t+1} - \mu\}\right] \text{ ------(2)}$$

And the auto-correlation function at lag k is given by-

$$\rho_k = \frac{\text{cov}(Z_t, Z_{t+k})}{\sqrt{V(Z_t) \times V(Z_{t+k})}}$$

$$=> \rho_k = \frac{E\left[\{Z_t - \mu\}\{Z_{t+1} - \mu\}\right]}{\sqrt{\sigma \times \sigma}} \quad \text{\{Using eq}^n\text{(a) and eq}^n\text{ (b) and eq}^n\text{ (2)\}}$$

$$=> \rho_k = \frac{E\left[\{Z_t - \mu\}\{Z_{t+1} - \mu\}\right]}{\sigma^2}$$

{ The above formula can be apply only for population values }

The sample estimate of R_k i.e. the sample auto-covariance function at lag k is given by-

$$c_k = \frac{1}{N} \sum_{t=1}^{N-k} \{u_t - \bar{u}\}\{u_{t+k} - \bar{u}\}$$

And the sample auto-correlation function at lag k is given by-

$$r_k = \frac{c_k}{c_0} \qquad \text{Where,} \qquad c_0 = \frac{1}{N}\sum_{t=1}^{N}\{u_t - \bar{u}\}^2$$

Calculation-

For lag 1, we construct the following table-

t	u_t	u_{t+1}	$\{u_t - \bar{u}\}$	$\{u_{t+1} - \bar{u}\}$	$\{u_t - \bar{u}\}\{u_{t+1} - \bar{u}\}$	$\{u_t - \bar{u}\}^2$
1	200	202	-2.757575758	-0.757575758	2.089072544	7.604224059
2	202	208	-0.757575758	5.242424242	-3.971533517	0.573921028
3	208	204	5.242424242	1.242424242	6.513314968	27.48301194
4	204	202	1.242424242	-0.757575758	-0.941230487	1.543617998
5	202	201	-0.757575758	-1.757575758	1.331496786	0.573921028
6	201	200	-1.757575758	-2.757575758	4.846648301	3.089072544
7	200	199	-2.757575758	-3.757575758	10.36179982	7.604224059
8	199	201	-3.757575758	-1.757575758	6.604224059	14.11937557
9	201	198	-1.757575758	-4.757575758	8.361799816	3.089072544
10	198	200	-4.757575758	-2.757575758	13.11937557	22.63452709
11	200	203	-2.757575758	0.242424242	-0.668503214	7.604224059
12	203	202	0.242424242	-0.757575758	-0.183654729	0.058769513
13	202	211	-0.757575758	8.242424242	-6.24426079	0.573921028
14	211	204	8.242424242	1.242424242	10.2405877	67.93755739
15	204	206	1.242424242	3.242424242	4.028466483	1.543617998
16	206	203	3.242424242	0.242424242	0.786042241	10.51331497
17	203	203	0.242424242	0.242424242	0.058769513	0.058769513
18	203	204	0.242424242	1.242424242	0.301193756	0.058769513
19	204	207	1.242424242	4.242424242	5.270890725	1.543617998
20	207	206	4.242424242	3.242424242	13.75573921	17.99816345
21	206	207	3.242424242	4.242424242	13.75573921	10.51331497

22	207	206	4.242424242	3.242424242	13.75573921	17.99816345
23	206	200	3.242424242	-2.757575758	-8.941230487	10.51331497
24	200	207	-2.757575758	4.242424242	-11.69880624	7.604224059
25	207	205	4.242424242	2.242424242	9.513314968	17.99816345
26	205	200	2.242424242	-2.757575758	-6.183654729	5.028466483
27	200	195	-2.757575758	-7.757575758	21.39210285	7.604224059
28	195	202	-7.757575758	-0.757575758	5.876951331	60.17998163
29	202	204	-0.757575758	1.242424242	-0.941230487	0.573921028
30	204	203	1.242424242	0.242424242	0.301193756	1.543617998
31	203	198	0.242424242	-4.757575758	-1.153351699	0.058769513
32	198	200	-4.757575758	-2.757575758	13.11937557	22.63452709
33	200		-2.757575758			7.604224059

From the above table we get-

N=33

$$\sum_{i=1}^{32}(u_t-\bar{u})(u_{t+1}-\bar{u})=124.456$$

$$\sum_{i=1}^{33}(u_t-\bar{u})^2=366.0606$$

Therefore, the sample auto-covariance function at lag 1 is denoted by c_1 and it is given by-

$$c_1=\frac{1}{N}\sum_{i=1}^{32}(u_t-\bar{u})(u_{t+1}-\bar{u})=\frac{1}{33}\times 124.456=3.771406$$

And, the sample variance term is denoted by c_0 and it is given by-

$$c_0=\frac{1}{N}\sum_{i=1}^{33}(u_t-\bar{u})^2=\frac{1}{33}\times 366.0606=11.09275$$

Therefore the sample auto-correlation function at lag 1 is given by-

$$r_1 = \frac{c_1}{c_0} = \frac{3.771406}{11.09275} = 0.339988$$

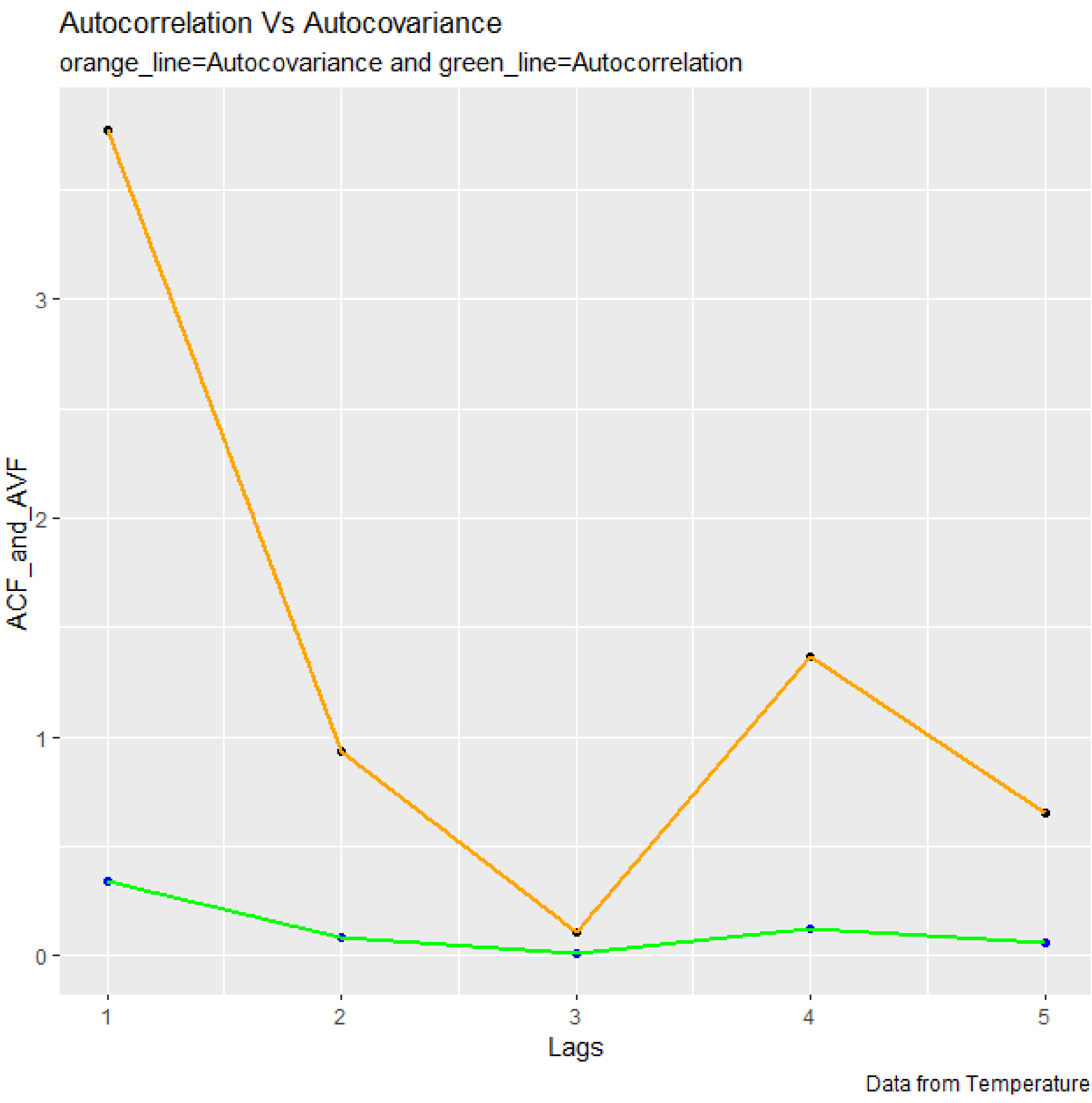
Programming in R

We can also find the sample Auto covariance and Auto correlation function by using the following R-program-

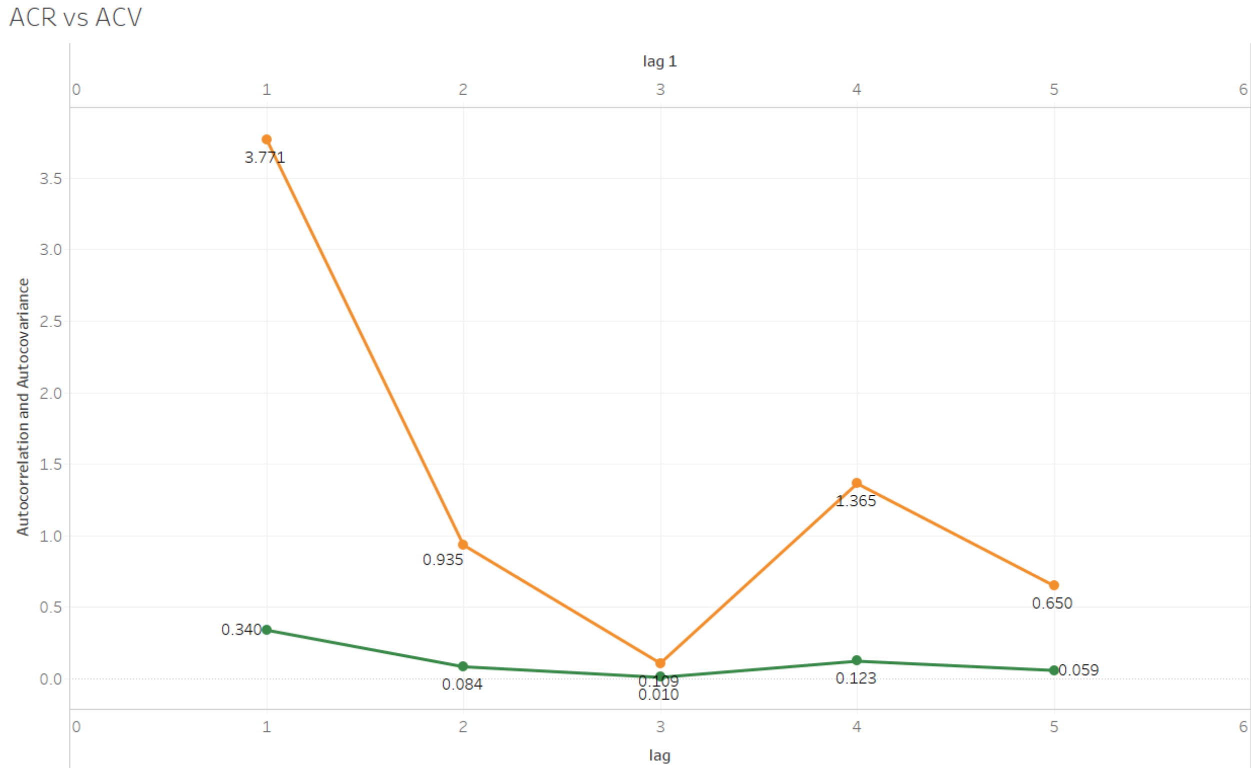
```
library('ggplot2')

ut=c(200, 202, 208, 204, 202, 201, 200, 199, 201, 198, 200, 203, 202, 211, 204, 206, 203, 203, 204, 207, 206, 207, 206, 200, 207, 205, 200, 195, 202, 204, 203, 198, 200)

N=length(ut)

N

Mean=mean(ut)

Mean

c0=var(ut)*((N-1)/N)

c0

c=mat.or.vec(5,1)

r=mat.or.vec(5,1)

for(i in 1:5){

  c[i]=(sum((ut[1:(33-i)]-Mean)*(ut[(i+1):33]-Mean)))/N

  r[i]=c[i]/c0

}

c

r
```

```
lag=c(1,2,3,4,5)

lag

Table_1 = data.frame(c,lag)

Table_1

View(Table_1)

Table_2 = data.frame(r,lag)

Table_2

View(Table_2)

ggp = ggplot(NULL, mapping = aes(x = lag,y = ACF_and_AVF)) +

        geom_point(data = Table_1, mapping = aes(x=lag,y=c), col = "black") + geom_line(data =
Table_1, mapping = aes(x=lag,y=c), col = "orange", size = 1) +

       geom_point(data = Table_2, mapping = aes(x=lag,y=r), col = "blue") + geom_line(data =
Table_2, mapping = aes(x=lag,y=r), col = "green", size = 1) +

labs(

  title = paste("Autocorrelation Vs Autocovariance"),

  subtitle = paste("orange_line=Autocovariance and green_line=Autocorrelation"),

  caption = "Data from Temperature",

  x = "Lags",

  y = "ACF_and_AVF"

 )

 ggp
```

Autocovariance vs Autocorrelation Graph using ggplot2-



Plotting the graph using Tableau



The trends of c and r for lag and lag 1. Color shows details about c and r. The data is filtered on r, which keeps all values.

Measure Names
c
r