QUALITY ENGINEERING

28/08/2019

General recommendations:

- avoid (if not required) theoretical introductions or explanations covered during the course;
- always state the assumptions, formulas/expressions and the final results (when using hypothesis tests provide the numerical value of the test statistic and the test conclusion in terms of p-value);
- show (qualitatively) all the plots.
- Duration: 2h (+ 10’ for manual data entry)

Instructions on how to access the software (Minitab and Excel) on the workstation:
Open the browser (Internet Explorer); Go on favourites and select “Polimi Virtual Desktop”; Enter your credentials. Do not use the link to Excel SW on the desktop.

Instructions on how to install the Solver in Excel:
Open Excel; go on “File\Options”; click Add-Ins, and then in the Manage box, select Excel Add-ins; Click Go; in the Add-Ins available box, select the Solver (Risolutore) and then click OK.

Exercise 1 (max score 14)

A manufacturing company executed a test campaign by measuring the force (in kN) in a stamping process for the production of 40 automotive components. The components were produced in the temporal order shown in the table. The time between one stamping process and the following one was 30 minutes.

component	F	component	F
1	2,862	21	3,402
2	3,697	22	2,295
3	2,355	23	2,566
4	3,002	24	3,495
5	2,901	25	3,384
6	3,257	26	2,599
7	2,911	27	2,771
8	2,948	28	2,377
9	3,517	29	3,918
10	2,712	30	3,418
11	2,617	31	2,130
12	3,782	32	3,115
13	3,373	33	3,522
14	2,372	34	3,190
15	2,281	35	2,508
16	3,777	36	2,481
17	3,445	37	3,749
18	2,714	38	3,767
19	2,347	39	2,281
20	3,680	40	2,550

a) Design a suitable control chart to determine if the stamping process was in-control or not based on the available measurements (design the chart in order to have an average time to signal equal to 50 h) – assume no assignable causes (when checking the assumption, refer to the LBQ test (first type error 5%) for the first 10 lags – show test results).

b) Considering that during the production of the 28th component, an improper placement of the material was recorded, how does the chart designed at point a) change?

Exercise 2 (max score 12)

A time series of individual measurements of a dimensional quality characteristic is available to monitor the stability of a milling process. The data recorded in temporal sequence are shown in the table below.

Obs.	X	Obs.	X
1	1,89	14	3,58
2	3,95	15	2,49
3	1,54	16	4,76
4	3,53	17	1,33
5	2,66	18	5,09
6	4,98	19	2,99
7	1,32	20	3,96
8	3,27	21	1,77
9	1,98	22	3,82
10	4,11	23	2,07
11	2,42	24	5,33
12	3,11	25	1,39
13	1,67	26	4,21

a) Check the assumptions of normality and randomness (use Bartlett's test – first type error 5% – at lag = 1)
b) Design a traditional I-MR control chart by neglecting any possible violation of the assumptions and comment results.
c) Design two control charts: one by applying gapping (one observation out of 2) and one by applying batching (batch size equal to 2) and comment the results; when needed use Bartlett's test – 5% first type error – at lag = 1.

Exercise 3 (max score 4)

A linear model $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$ is fitted to a time series consisting of n=150 observations and the following result was obtained:

Term	Coef	SE Coef	T-value	P-value
β_0	0.0121	0.0585	0.21	0.838
β_1	-0.8692	0.0947	-9.18	0.000
β_2	-0.536	0.142	-3.78	0.001

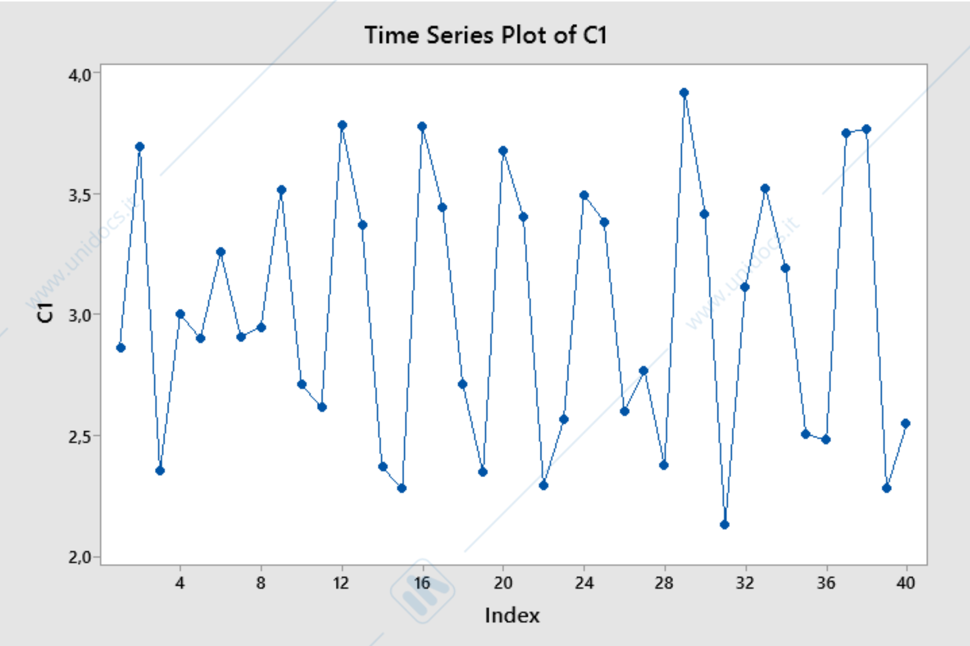
Estimate the familywise 95% confidence interval for the significant coefficients of the model.

SOLUTIONS

Exercise 1

a)

The time series of the stamping force exhibits a non-random pattern, possibly characterized by a negative autocorrelation:



Runs test:

Runs Test: C1

Descriptive Statistics

		Number of Observations	
N	K	≤ K	> K
40	3,00169	21	19

K = sample mean

Test

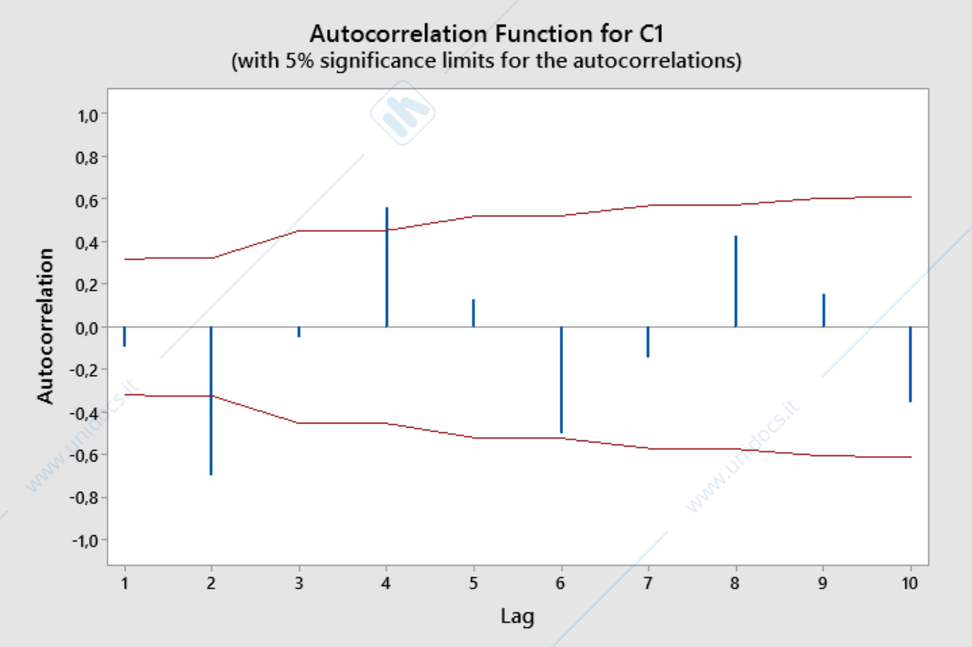
Null hypothesis H_0: The order of the data is random
Alternative hypothesis H_1: The order of the data is not random

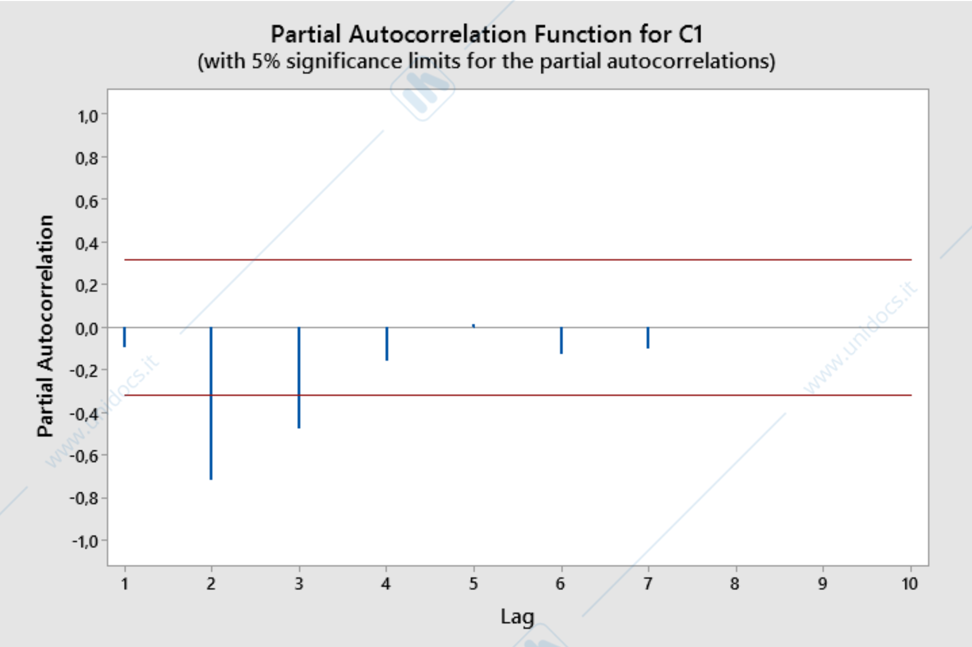
Number of Runs		
Observed	Expected	P-Value
23	20,95	0,510

LBQ test (up to lag 10):

Test statistic: LBQ = 68.98.
Rejection region at 95% confidence: 18.31
Since LBQ = 68.98>18.31, the null assumption of non autocorrelated data shall be rejected.

The sample autocorrelation and partial autocorrelation functions are reported below. Although there is no "clean" pattern to identify a suitable ARIMA model, there is a large partial autocorrelation up to lag 3 (by looking to the PACF), which is responsible for the pattern shown in the time series plot.





One suitable model for this time series is an AR(3). The non-random pattern of the stamping force can not be properly fitted by means of neither autoregressive models of lower order nor AM models. It is also possible to verify that ARMA models with order ≤ 3 do not lead to random residuals.

Regression Analysis: C1 versus ar1; ar2; ar3

Method

Rows unused 3

Regression Equation

C1 = 8,870 - 0,537 ar1 - 0,8692 ar2 - 0,536 ar3

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	8,870	0,913	9,72	0,000	
ar1	-0,537	0,138	-3,89	0,000	2,33
ar2	-0,8692	0,0947	-9,18	0,000	1,09
ar3	-0,536	0,142	-3,78	0,001	2,31

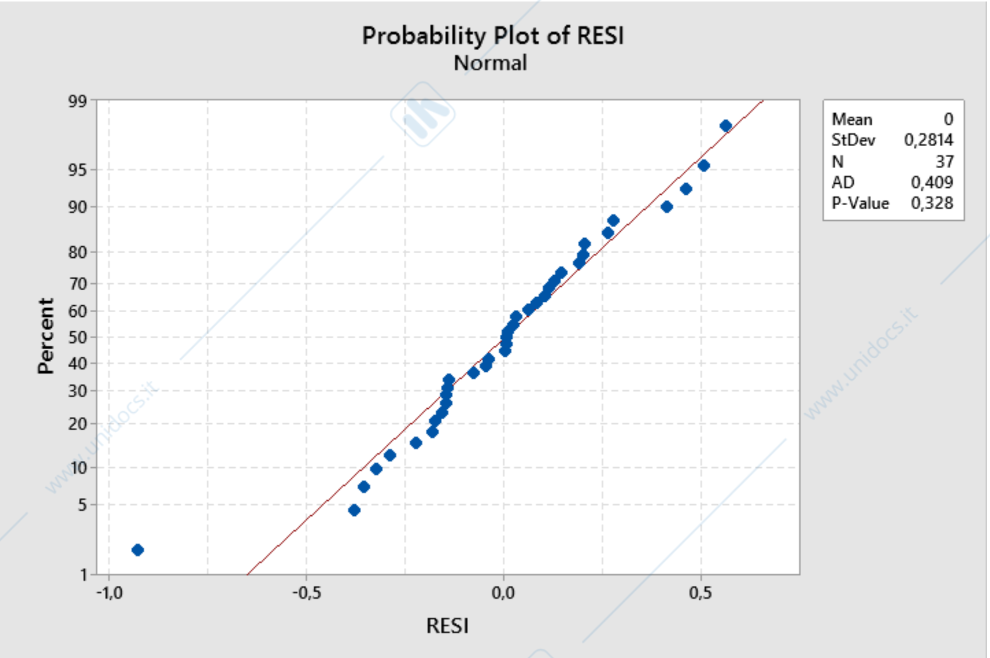
Model Summary

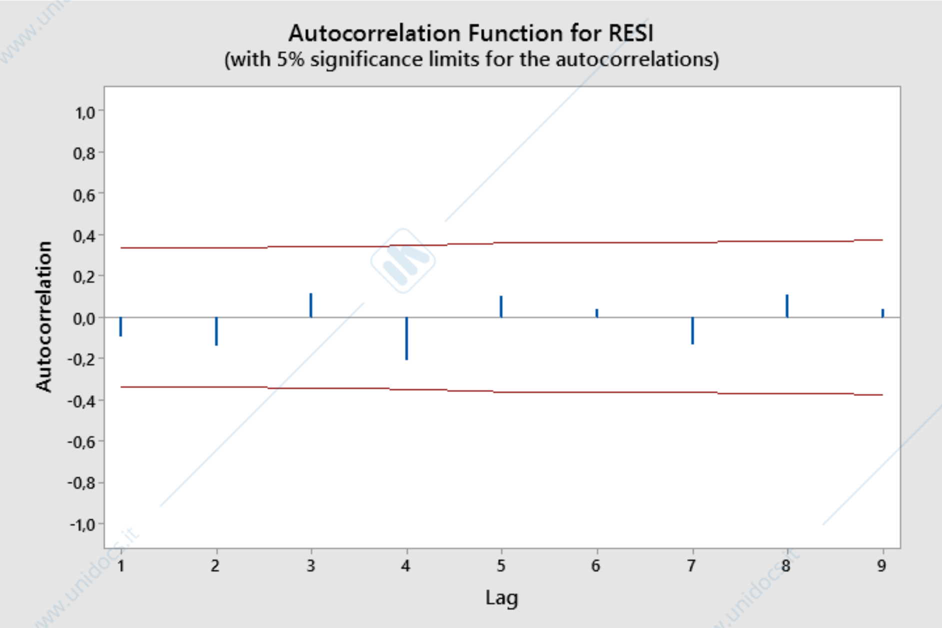
S	R-sq	R-sq(adj)	R-sq(pred)
0,293873	72,42%	69,92%	67,03%

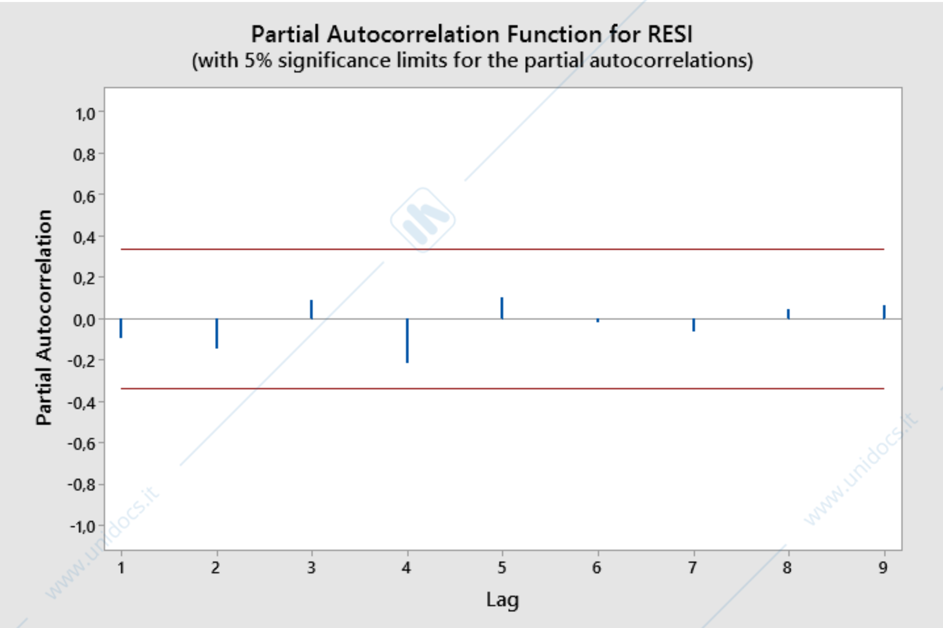
Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	3	7,485	2,49503	28,89	0,000
ar1	1	1,304	1,30442	15,10	0,000
ar2	1	7,273	7,27265	84,21	0,000
ar3	1	1,236	1,23577	14,31	0,001
Error	33	2,850	0,08636		
Total	36	10,335			

The residuals are normal and independent:







Runs test:

Runs Test: RESI

Descriptive Statistics

		Number of Observations	
N	K	≤ K	> K
37	0	16	21

K = sample mean

Test

Null hypothesis H_0: The order of the data is random
Alternative hypothesis H_1: The order of the data is not random

Number of Runs		
Observed	Expected	P-Value

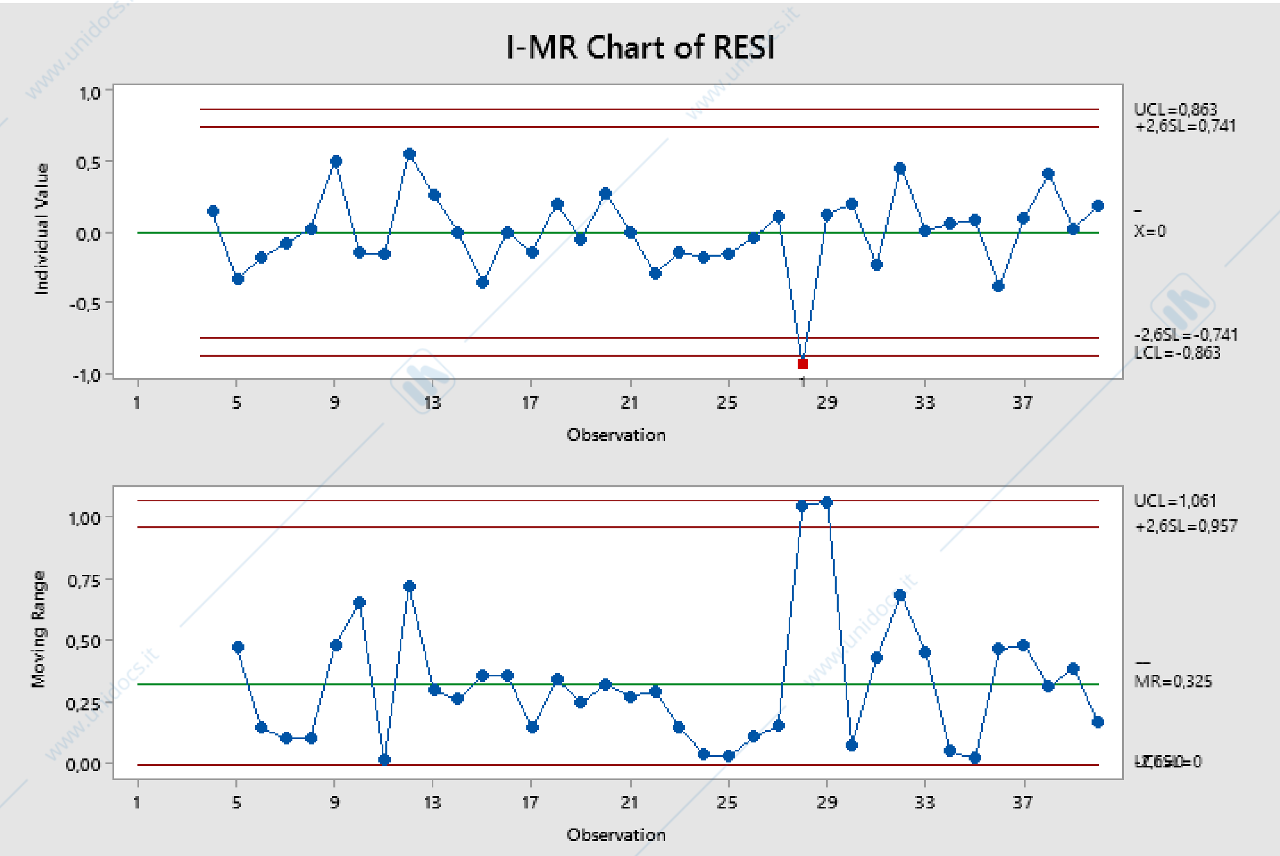
19	19,16	0,956

LBQ test (lag = 10):

Test statistic: LBQ =5.56.
Rejection region at alpha 5%: 18.31
Since LBQ = 5.56<18.31, the null assumption of non autocorrelated data shall not be rejected.

The average time to signal is ATS = ARL*t = 50 h, where t is the time between one measurement and the following one. In this case t = 30 min, thus ARL = 100, and hence $k = z_{\alpha/2} = 2{,}576$.

The resulting control chart is the following:



There is an out of control on the I chart and two out of control on the MR chart. Assuming no assignable causes, the control chart design is over. Two

b)

Considering the additional information for component 28, it is possible to include a dummy variable into the AR(3) model as follows:

Regression Analysis: C1 versus ar1; ar2; ar3; dummy

Method

Categorical predictor coding	(1; 0)
Rows unused	3

Regression Equation

dummy			
0	C1	=	8,943 - 0,542 ar1 - 0,9087 ar2 - 0,507 ar3
1	C1	=	7,956 - 0,542 ar1 - 0,9087 ar2 - 0,507 ar3

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	8,943	0,763	11,72	0,000	
ar1	-0,542	0,116	-4,69	0,000	2,33
ar2	-0,9087	0,0798	-11,38	0,000	1,11
ar3	-0,507	0,119	-4,27	0,000	2,32
dummy					
1	-0,987	0,253	-3,90	0,000	1,03

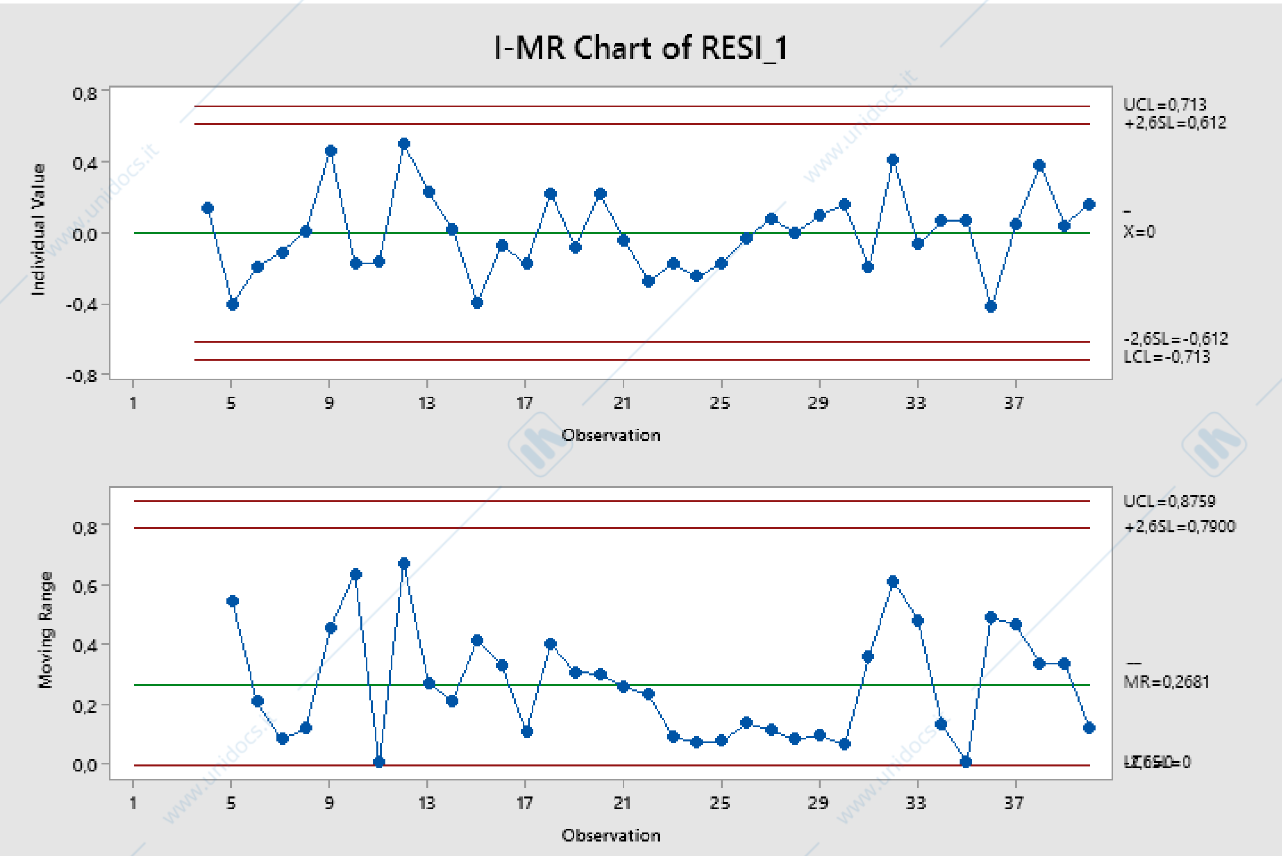
Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0,245660	81,31%	78,98%	*

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	4	8,4039	2,10096	34,81	0,000
ar1	1	1,3292	1,32916	22,02	0,000
ar2	1	7,8208	7,82078	129,59	0,000
ar3	1	1,0985	1,09852	18,20	0,000
dummy	1	0,9188	0,91876	15,22	0,000
Error	32	1,9312	0,06035		
Total	36	10,3350			

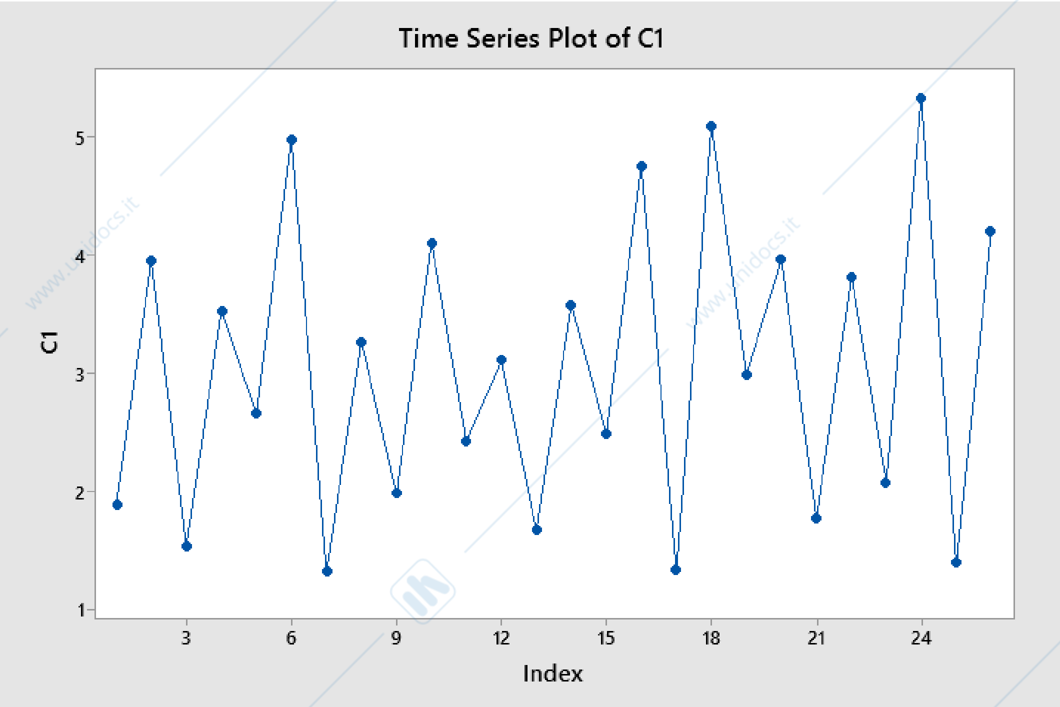
All the terms in the model are significant and the residuals are normal and independent. The resulting control chart is the following, where no out of control is signaled. The process is in-control. An unusual "pattern" is clear in the interval 22-30 (in terms of individual and range) but no alarm is issued and the process is deemed in-control



Exercise 2

a)

The time series of the measured quality characteristics exhibits a pattern that is possibly due to a negative auto-correlation at lag 1:



Runs test:

Runs Test: C1

Descriptive Statistics

		Number of Observations	
N	K	≤ K	> K
26	3,04692	13	13

K = sample mean

Test

Null hypothesis H_0: The order of the data is random
Alternative hypothesis H_1: The order of the data is not random

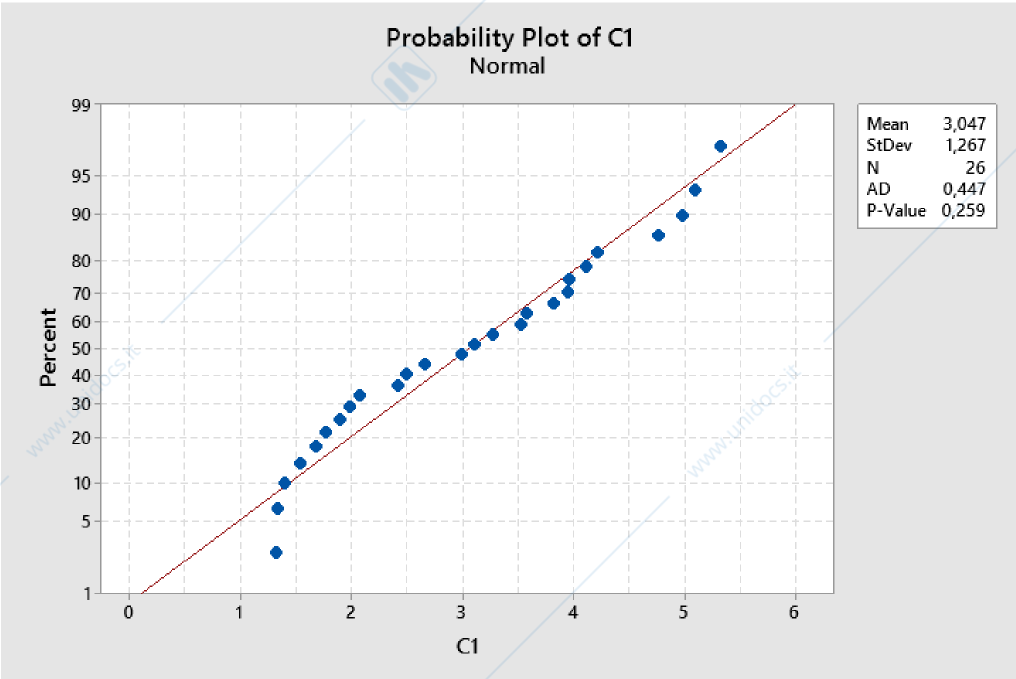
Number of Runs		
Observed	Expected	P-Value
26	14,00	0,000

Bartlett's test (alpha=5%) at lag = 1:

Test statistic at lag 1: r1 = -0.733018
The critical region is: $|r1| > \frac{z_{\alpha/2}}{\sqrt{n}} = 0{,}384$, where n=26 and $z_{\alpha/2} = 1{,}96$

Since $|r1|=0.733018>0.384$, the null hypothesis of lack of autocorrelation at lag 1 is rejected.

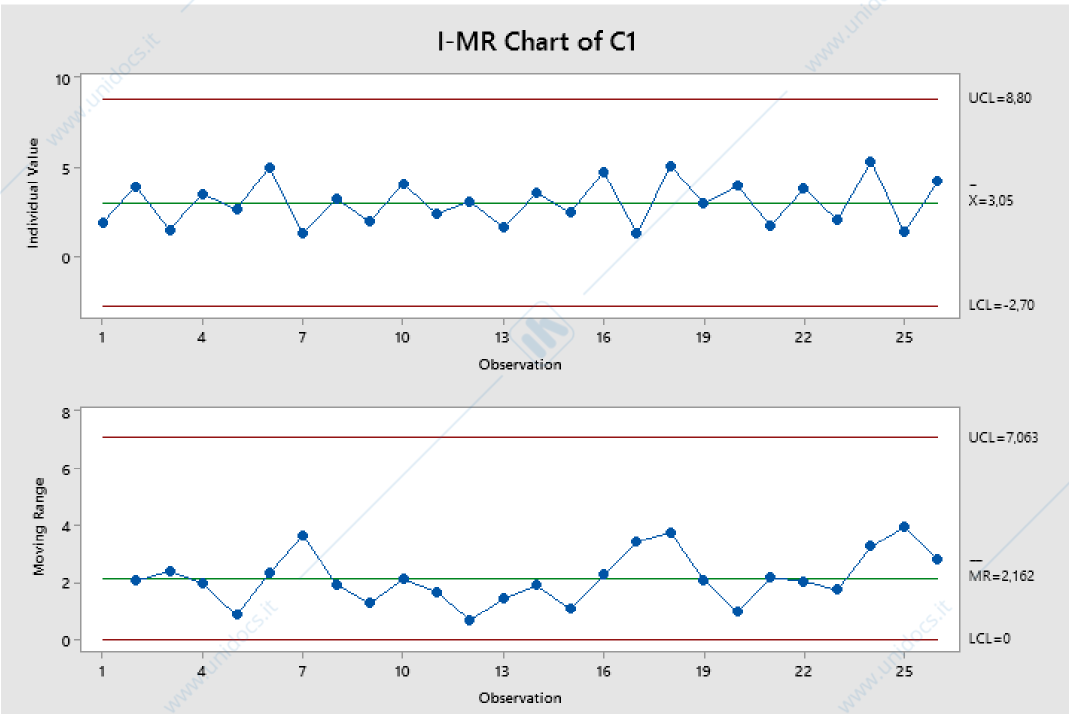
Normality test:



Data are normal but autocorrelated.

b)

If the violation of assumptions is neglected the following control chart is obtained:

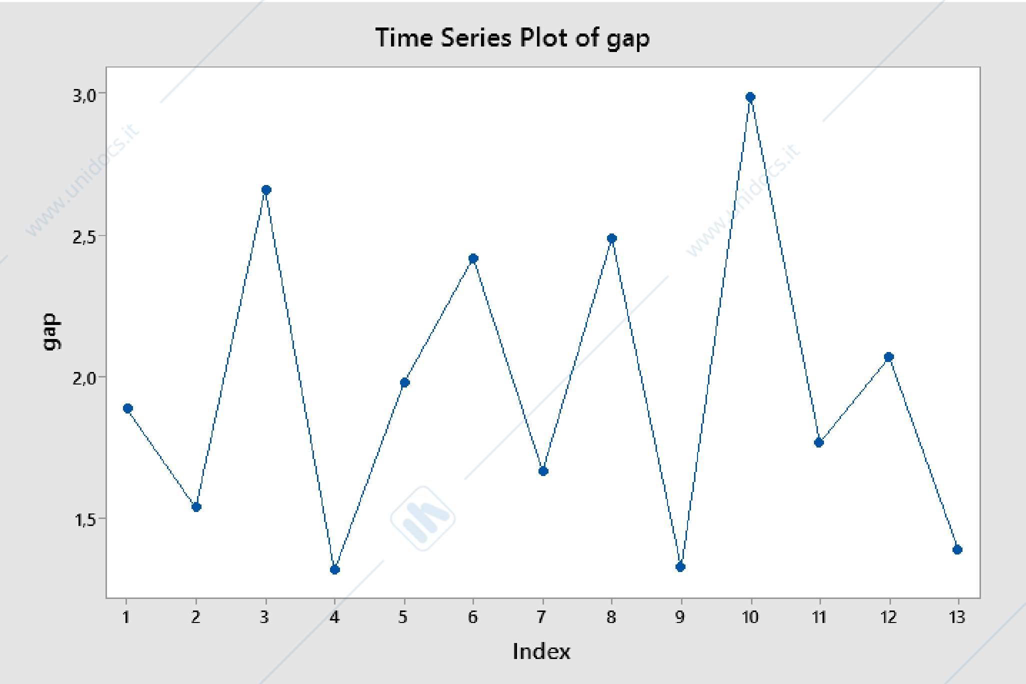


Although no out-of-control is signaled, there is an evident out-of-control pattern in the I chart that combines a non-random pattern with a hugging effect.

c)

Gapping

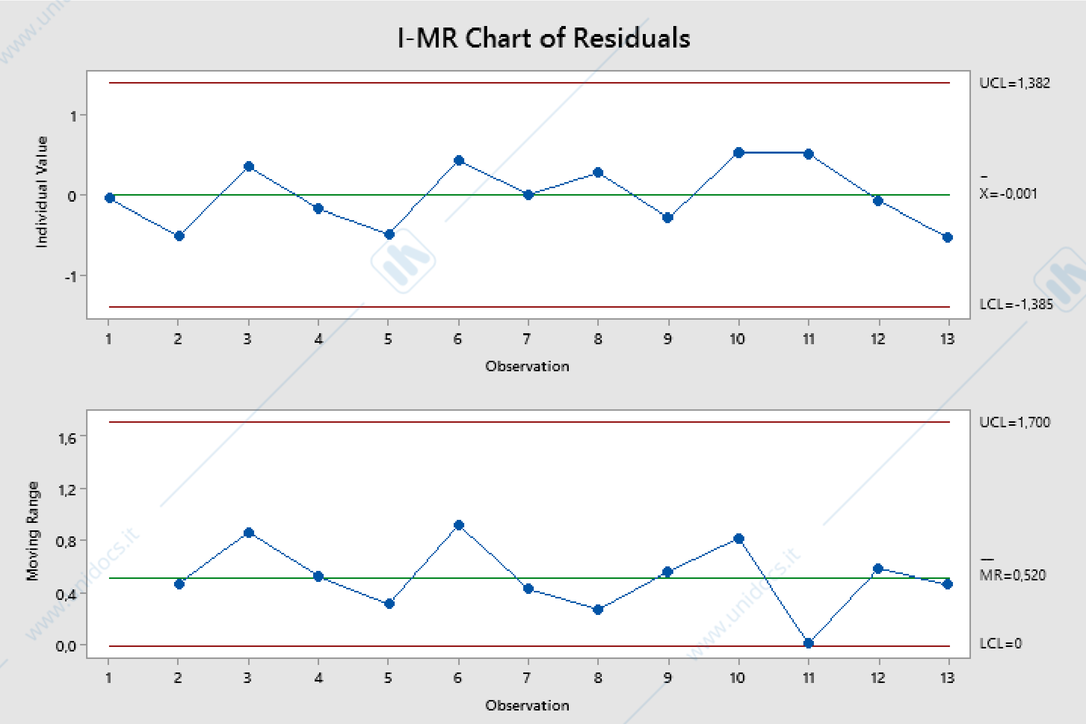
The time series of data after applying gapping is the following:



The p-values of runs test, Bartlett's tests (lag 1) and normality test are the following:

	p-value
Runs test	0.039
Bartlett's test	0.009553
Normality test	0.539

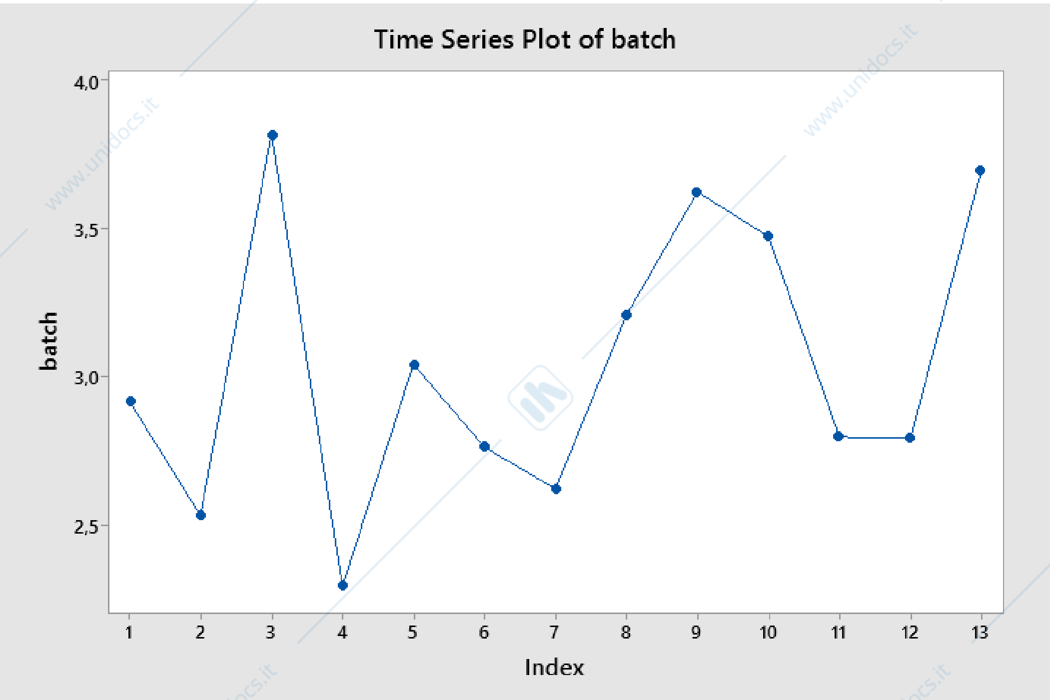
The gapping operation is not sufficient to deal with the violation of assumptions, and it also leads to a loss of information regarding the underlying temporal pattern in the original time series.
It must be also pointed out that the overall sample size after gapping is now very small (only 13). New data should be observed before defining the final monitoring approach.
Due to the lack of randomness of the gapped time series, a traditional I-MR control chart is not applicable, and the gapping operation may be considered not suitable in this case.
Alternatively, an AR(1) model fitted to the gap statistic ($Y(t) = 3.428 - 0.727\ Y(t-1)$) yields normal (p-value = 0.355) and independent (Bartlett test at lag 1 p-value = 0.2772) residuals: the resulting I-MR control chart is the following:



Despite the loss of information entailed by the gapping operation, the process is in-control even if the hugging seems to remain.

Batching

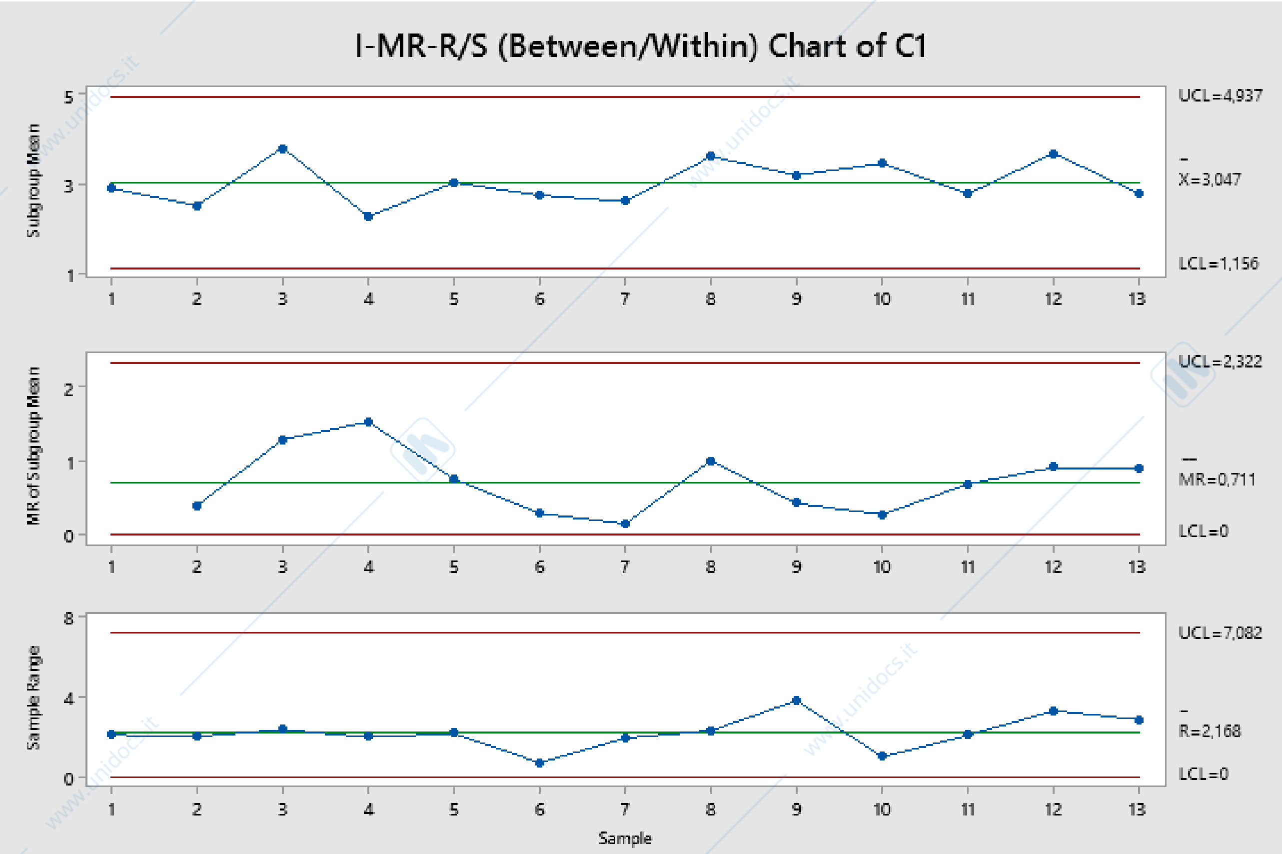
The time series of data after applying batching is the following:



The p-values of runs test, Bartlett's tests (lag 1) and normality test are the following:

	p-value
Runs test	0.925
Bartlett's test	0.04612
Normality test	0.476

The same conclusions for the gapping approach can be applied in this case, too.
According to the Bartlett test the batched time series is barely random, but it appears that the batching operation is not fully suitable to get rid of the original violation of assumptions.
Accepting the border-line randomness of the batched time-series, it is possible to design an I-R-MR control chart to monitor the mean of the batches, the variability within the batches and the variability between the batches:



The process is in-control, although some hugging effect is still present, especially in the sample range chart for the first batches.

In this case, both the batching and gapping operations reduce the data set to a very small sample size. New data should be taken to arrive to the final conclusion. Additional check on why the data have the negative autocorrelation pattern should be also interesting for the plant manager.

Exercise 3

The confidence interval for linear model coefficients is defined as follows:

$$\hat{\beta}_i - t_{\frac{\alpha}{2},n-p} se(\hat{\beta}_i) \leq \beta_i \leq \hat{\beta}_i + t_{\frac{\alpha}{2},n-p} se(\hat{\beta}_i).$$

In this case, the constant term is not significant, whereas both β_1 and β_2 are significant. The Bonferroni's correction should be used, such as, being $\alpha' = 0.05$, the individual confidence interval is designed with $\alpha = \frac{\alpha'}{2} = 0.025$.

Being:

- n=150
- p=3
- $t_{\frac{\alpha}{2},n-p}$=2.265

The 95% familywise confidence intervals are:

$$-1.0847 \leq \beta_1 \leq -0.6547$$

$$-0.6547 \leq \beta_2 \leq -0.21437$$