

АУТОРЕГРЕСИВНИ МОДЕЛ

-ПРИМЕР-

На основу података о висини прихода у милионима динара оствареним у периоду од 2005-2018. године дефинисати ауторегресивне моделе првог, другог и трећег реда и одабрати модел који највише одговара датим подацима уз ниво поверења од 95%.

Година	Приходи
2005	18
2006	18,5
2007	18,9
2008	18,8
2009	19,8
2010	20,5
2011	20,1
2012	19,6
2013	21
2014	21,9
2015	23,1
2016	24,1
2017	28,9
2018	31,9

РЕШЕЊЕ:

- Прво се дефинишу Lag 1, Lag 2 и Lag 3, копирањем вредности из колоне „Приходи“.

Табела 1.

Година	Приходи	Lag 1	Lag 2	Lag 3
2005	18	#N/A	#N/A	#N/A
2006	18,5	18	#N/A	#N/A
2007	18,9	18,5	18	#N/A
2008	18,8	18,9	18,5	18
2009	19,8	18,8	18,9	18,5
2010	20,5	19,8	18,8	18,9
2011	20,1	20,5	19,8	18,8
2012	19,6	20,1	20,5	19,8
2013	21	19,6	20,1	20,5
2014	21,9	21	19,6	20,1

Not Available/
Подаци нису
расположиви

2015	23,1	21,9	21	19,6
2016	24,1	23,1	21,9	21
2017	28,9	24,1	23,1	21,9
2018	31,9	28,9	24,1	23,1

- Затим се одређује **ауторегресивни модел трећег реда** који указује на везу између вредности које су удаљене 3 периода по следећем поступку:

Data→Data Analysis→Regression

Input Y range: обухватити следећу колону:

Табела 2.

Година	Приходи	Lag 1	Lag 2	Lag 3
2005	18	#N/A	#N/A	#N/A
2006	18,5	18	#N/A	#N/A
2007	18,9	18,5	18	#N/A
2008	18,8	18,9	18,5	18
2009	19,8	18,8	18,9	18,5
2010	20,5	19,8	18,8	18,9
2011	20,1	20,5	19,8	18,8
2012	19,6	20,1	20,5	19,8
2013	21	19,6	20,1	20,5
2014	21,9	21	19,6	20,1
2015	23,1	21,9	21	19,6
2016	24,1	23,1	21,9	21
2017	28,9	24,1	23,1	21,9
2018	31,9	28,9	24,1	23,1

Input X range: обухватити следеће колоне:

Табела 3.

Година	Приходи	Lag 1	Lag 2	Lag 3
2005	18	#N/A	#N/A	#N/A
2006	18,5	18	#N/A	#N/A
2007	18,9	18,5	18	#N/A
2008	18,8	18,9	18,5	18
2009	19,8	18,8	18,9	18,5
2010	20,5	19,8	18,8	18,9
2011	20,1	20,5	19,8	18,8
2012	19,6	20,1	20,5	19,8

2013	21	19,6	20,1	20,5
2014	21,9	21	19,6	20,1
2015	23,1	21,9	21	19,6
2016	24,1	23,1	21,9	21
2017	28,9	24,1	23,1	21,9
2018	31,9	28,9	24,1	23,1

Confidence level: 95% (ако је у поставци задатка 95%, односно 99%)

New Worksheet Ply: AR-3 (произвољно одабрати назив за нови лист, али је препоручљиво да има везе са оним што се рачуна)

OK

Од добијених резултата на новом радном листу посматрају се следећи:

Табела 4.

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	-15,10237588	6,39201786	-2,362693003	0,050142246
X Variable 1	0,884865495	0,357122052	2,477767729	0,042348344
X Variable 2	0,329428851	0,747620763	0,440636306	0,672763112
X Variable 3	0,598677216	0,718713568	0,832984437	0,432352284

На основу општег обрасца за оцењену једначину $\hat{Y}_i = a_0 + a_1Y_{i-1} + a_2Y_{i-2} + a_3Y_{i-3}$ дефинише се **ауторегресивна једначина трећег реда** која гласи:

$$\hat{Y}_i = -15,1024 + 0,8849Y_{i-1} + 0,3294Y_{i-2} + 0,5987Y_{i-3}$$

чије се вредности читавају из Табеле 4.

Тестира се нулта хипотеза за параметар највишег реда, односно трећег реда:

$$H_0: A_3 = 0$$

против алтернативне:

$$H_1: A_3 \neq 0$$

За вредности t теста узима се вредност коефицијента $a_3 = 0,5987$ и стандардне грешке $S_{a_3} = 0,7187$ чије се вредности такође налазе у Табели 4:

$$t = \frac{a_3 - A_3}{S_{a_3}} = \frac{0,5987 - 0}{0,7187} = 0,833$$

$$\alpha = 0,05$$

број степени слободе $n - 2p - 1 = 14 - 2 * 3 - 1 = 7$

$$t_{\frac{\alpha}{2}, 7} = t_{0,025; 7} = \pm 2,3646$$

Зато што је:

$$-2,3646 < t = 0,833 < +2,3646$$

и

$$p = 0,4324 > 0,05 \text{ (очитава се из Табеле 4.)}$$

не одбацује се нулта хипотеза и закључујемо да параметар трећег реда ауторегресивног модела није статистички значајан и да се може избрисати из модела.

- Затим се одређује **ауторегресивни модел другог реда** који указује на везу између вредности које су удаљене 2 периода по следећем поступку:

Data→Data Analysis→Regression

Input Y range: обухватити следећу колону:

Табела 5.

Година	Приходи	Lag 1	Lag 2	Lag 3
2005	18	#N/A	#N/A	#N/A
2006	18,5	18	#N/A	#N/A
2007	18,9	18,5	18	#N/A
2008	18,8	18,9	18,5	18
2009	19,8	18,8	18,9	18,5
2010	20,5	19,8	18,8	18,9
2011	20,1	20,5	19,8	18,8
2012	19,6	20,1	20,5	19,8
2013	21	19,6	20,1	20,5
2014	21,9	21	19,6	20,1
2015	23,1	21,9	21	19,6
2016	24,1	23,1	21,9	21
2017	28,9	24,1	23,1	21,9
2018	31,9	28,9	24,1	23,1

Input X range: обухватити следеће колоне:

Табела 6.

Година	Приходи	Lag 1	Lag 2	Lag 3
2005	18	#N/A	#N/A	#N/A
2006	18,5	18	#N/A	#N/A
2007	18,9	18,5	18	#N/A
2008	18,8	18,9	18,5	18
2009	19,8	18,8	18,9	18,5
2010	20,5	19,8	18,8	18,9
2011	20,1	20,5	19,8	18,8
2012	19,6	20,1	20,5	19,8
2013	21	19,6	20,1	20,5
2014	21,9	21	19,6	20,1
2015	23,1	21,9	21	19,6
2016	24,1	23,1	21,9	21
2017	28,9	24,1	23,1	21,9
2018	31,9	28,9	24,1	23,1

Confidence level: 95% (ако је у поставци задатка 95%, односно 99%)

New Worksheet Ply: AR-2 (произвољно одабрати назив за нови лист, али је препоручљиво да има везе са оним што се рачуна)

OK

Од добијених резултата на новом радном листу посматрају се следећи:

Табела 7.

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	-10,75793671	4,645857357	-2,315597722	0,045814793
X Variable 1	0,990312318	0,324640261	3,050491383	0,013781945
X Variable 2	0,593399661	0,511245939	1,160693151	0,275627345

На основу општег обрасца за оцењену једначину $\hat{Y}_i = a_0 + a_1Y_{i-1} + a_2Y_{i-2}$ дефинише се ауторегресивна једначина другог реда која гласи:

$$\hat{Y}_i = -10,7579 + 0,9903Y_{i-1} + 0,5934Y_{i-2}$$

чије се вредности читавају из Табеле 7.

Тестира се нулта хипотеза за параметар највишег реда, односно другог реда:

$$H_0: A_2 = 0$$

против алтернативне:

$$H_1: A_2 \neq 0$$

За вредности t теста узима се вредност коефицијента $a_2 = 0,5934$ и стандардне грешке $S_{a_2} = 0,5112$ чије се вредности такође налазе у Табели 7:

$$t = \frac{a_2 - A_2}{S_{a_2}} = \frac{0,5934 - 0}{0,5112} = 1,1607$$

$$\alpha = 0,05$$

број степени слободe $n - 2p - 1 = 14 - 2 * 2 - 1 = 9$

$$t_{\frac{\alpha}{2}, 9} = t_{0,025; 9} = \pm 2,2622$$

Зато што је:

$$-2,2622 < t = 1,1607 < +2,2622$$

и

$$p = 0,2765 > 0,05 \text{ (очитава се из Табеле 7.)}$$

не одбацује се нулта хипотеза и закључујемо да параметар другог реда ауторегресивног модела није статистички значајан и да се може избрисати из модела.

- Затим се одређује **ауторегресивни модел првог реда** који указује на везу између узастопних вредности по следећем поступку:

Data→Data Analysis→Regression

Input Y range: обухватити следећу колону:

Табела 8.

Година	Приходи	Lag 1	Lag 2	Lag 3
2005	18	#N/A	#N/A	#N/A
2006	18,5	18	#N/A	#N/A
2007	18,9	18,5	18	#N/A
2008	18,8	18,9	18,5	18
2009	19,8	18,8	18,9	18,5
2010	20,5	19,8	18,8	18,9

2011	20,1	20,5	19,8	18,8
2012	19,6	20,1	20,5	19,8
2013	21	19,6	20,1	20,5
2014	21,9	21	19,6	20,1
2015	23,1	21,9	21	19,6
2016	24,1	23,1	21,9	21
2017	28,9	24,1	23,1	21,9
2018	31,9	28,9	24,1	23,1

Input X range: обухватити следећу колону:

Табела 9.

Година	Приходи	Lag 1	Lag 2	Lag 3
2005	18	#N/A	#N/A	#N/A
2006	18,5	18	#N/A	#N/A
2007	18,9	18,5	18	#N/A
2008	18,8	18,9	18,5	18
2009	19,8	18,8	18,9	18,5
2010	20,5	19,8	18,8	18,9
2011	20,1	20,5	19,8	18,8
2012	19,6	20,1	20,5	19,8
2013	21	19,6	20,1	20,5
2014	21,9	21	19,6	20,1
2015	23,1	21,9	21	19,6
2016	24,1	23,1	21,9	21
2017	28,9	24,1	23,1	21,9
2018	31,9	28,9	24,1	23,1

Confidence level: 95% (ако је у поставци задатка 95%, односно 99%)

New Worksheet Ply: AR-1 (произвољно одабрати назив за нови лист, али је препоручљиво да има везе са оним што се рачуна)

OK

Од добијених резултата на новом радном листу посматрају се следећи:

Табела 10.

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	-5,838372834	2,257434903	-2,586286243	0,025305187

X Variable 1	1,328692704	0,106438338	12,48321549	7,74478E-08
--------------	-------------	-------------	-------------	-------------

На основу општег обрасца за оцењену једначину $\hat{Y}_i = a_0 + a_1 Y_{i-1}$ дефинише се **ауторегресивна једначина првог реда** која гласи:

$$\hat{Y}_i = -5,8384 + 1,3287Y_{i-1}$$

чије се вредности читавају из Табеле 10.

Тестира се нулта хипотеза за параметар највишег реда, односно првог реда:

$$H_0: A_1 = 0$$

против алтернативне:

$$H_1: A_1 \neq 0$$

За вредности t теста узима се вредност коефицијента $a_1 = 1,3287$ и стандардне грешке $S_{a_1} = 0,1064$ чије се вредности такође налазе у Табели 10:

$$t = \frac{a_1 - A_1}{S_{a_1}} = \frac{1,3287 - 0}{0,1064} = 12,4832$$

$$\alpha = 0,05$$

број степени слободe $n - 2p - 1 = 14 - 2 * 1 - 1 = 11$

$$t_{\frac{\alpha}{2}, 11} = t_{0,025, 11} = \pm 2,2010$$

Зато што је:

$$t = 12,4832 > 2,2010$$

и

$$p = 0,000 < 0,05 \text{ (читава се из Табеле 10.)}$$

одбацује се нулта хипотеза и закључујемо да је параметар првог реда ауторегресивног модела статистички значајан и да треба да остане у моделу.

- На основу оцењеног ауторегресивног модела првог реда израчунавају се предвиђене вредности прихода и врши се предвиђање за 2019. годину:

Табела 11.

Година	Приходи	Предвиђене вредности AR-1
2005	18	-

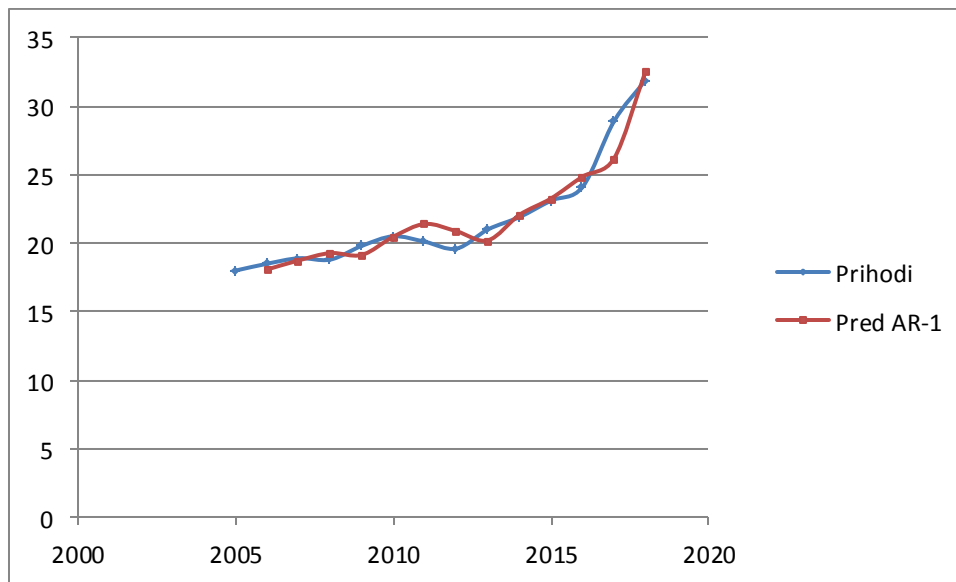
2006	18,5	18,0782
2007	18,9	18,74255
2008	18,8	19,27403
2009	19,8	19,14116
2010	20,5	20,46986
2011	20,1	21,39995
2012	19,6	20,86847
2013	21	20,20412
2014	21,9	22,0643
2015	23,1	23,26013
2016	24,1	24,85457
2017	28,9	26,18327
2018	31,9	32,56103
2019		36,54713

$$\hat{Y}_i = -5,8384 + 1,3287 \cdot 18 = 18,0782\dots$$

На крају се означе све три колоне из Табеле 11. и следи:

Insert→Scatter→Scatter with Smooth Lines and Markers

да би се добио следећи графички приказ:



Слика 1.