АНАЛИЗА ВРЕМЕНСКИХ СЕРИЈА

Низ података са природним, хронолошким редоследом назива се *временска серија*. Овако уређени подаци срећу се у разним областима. На пример, у економији, као месечни индекси цена на мало, вредност акција на берзи током одређеног временског периода, итд.; у геофизици, сеизмограми којима се региструју потреси земљине коре; у медицини, кардиограми и сл.

Статистичка анализа временских серија заснована је на поставкама теорије вероватноће и статистике. У анализи временских серија претпоставака о независности опсервација као код случајног узорка не важи. Постојање природног, логичног поретка променљивих, односно чланова временске серије, условљава специфичност анализе временских серија.

Најважнији проблем у анализи временских серија јесте *предвиђање*, односно посматрање промена током времена и пројектовање будућег кретања. Као такво, предвиђање се може користити у профитним и непрофитним организација, а ради предвиђања тражње за производима, потреба потрошача, нивоа незапослености, инфлације, уписа студената, итд.

У анализи временских серија присутне су и врло једноставне, али и сложеније методе које су засноване на теорији случајних или стохастичких процеса, пошто временска серија и није ништа друго него делимична реализација бесконачног низа случајних променљивих, односно случајног процеса. Математичка формулација начина генерисања серије назива се *модел временске серије*. Полазна претпоставка је да ће фактори који су утицали на појаву у прошлости и који делују на појаву у садашњости, на исти начин утицати на појаву у будућности.

За анализу је често од интереса издвајање одређене компоненте временске серије. Економска временска серија се може схватити као збир или производ разних компоненти, од којих се најчешће користе следеће:

- *тренд* - представља дугорочну тенденцију развоја серије;
- *циклуси* - представљају осцилације са променљивим периодом који је дужи од 12 месеци и зато се често та компонента не раздваја од тренда;
- *сезонска компонента* - има периодику од 12 месеци и показује утицај појединих месеци у години;
- *ирегуларна или случајна компонента* - ову компоненту чине неправилне осцилације, које у економској серији могу бити последице штрајка или неког другог узрока.

ТРЕНД

Под трендом се подразумева дугорочна развојна тенденција варирања посматране појаве током времена. Код испитивања тренда полази се од претпоставке да на развој појаве делују одређени фактори. Једна група фактора делује у одређеном правцу, док друга група делује на скретање тока појаве навише или наниже у односу на устаљени правац.

Временска серија x_t од n чланова увек се може тачно представити неком функцијом од t:

$$x_t = f(t)$$

За $n = 2$, серија би имала само два члана који би се тачно представили правом кроз две тачке, односно функцијом $f(t)$ која је у том случају једначина праве. У општем случају, за серију од n чланова, функција која би је тачно представила имала би n параметара.

Сврха представљања функције моделом као што је $x_t = f(t)$ јесте могућност прогнозирања њених будућих чланова. Прогноза члана x_{t+1} била би $f(t+1)$ итд. Даље, модел $x_t = f(t)$ модификује се додавањем случајне компоненте ε_t која у себи садржи грешку мерења, односно сва одступања вредности временске серије од детерминисане функције $f(t)$. Одговарајући модел је:

$$x_t = f(t) + \varepsilon_t \qquad t = 1,2,\ldots,n$$

где случајна компонента ε_t представља низ независних случајних промењивих са очекиваном вредности једнакој нули.

Систематска компонента $f(t)$ је познатог аналитичког облика, са непознатим параметрима, који се, као код регресије, оцењују методом најмањих квадрата. Због тога се модели типа $x_t = f(t) + \varepsilon_t$ називају регресиони модели. Када је $f(t)$ линеарна функција, онда модел може да представи серију са доминантним линеарним трендом.

ЛИНЕАРНИ ТРЕНД

Када временска серија показује приближно праволинијску тенденцију, онда кажемо да је присутан *линеарни тренд*. Код линеарног тренда, варијације временске серије изравнавају се линијом тренда која показује просечно кретање појаве на дуги рок.

Оцењени модел

Теоријска функција линеарног тренда гласи:

$$y_i = \beta_0 + \beta_1 \cdot x_i$$

а оцењена:

$$\hat{y}_i = b_0 + b_1 \cdot x_i$$

где је:

$\hat{y}_i$ – оцена просечних вредности посматране појаве;

x_i – „кодиране“ године;

b_0 и b_1 – оцене параметара β_0 и β_1.

Параметре b_0 и b_1 можемо одредити на основу следећих образаца:

$$b_0 = \bar{y} = \frac{\sum y_i}{n}$$

$$b_1 = \frac{\sum x_i y_i}{\sum x_i^2}$$

Вредност параметра b_0 показује просечну вредност посматране појаве у одређеном временском периоду, док коефицијент b_1 показује средњи апсолутни пораст. У зависности од предзнака испред коефицијента b_1, линија тренда биће растућа или опадајућа.

Код серије са непарним бројем података, средишњи временски период (година, месец, дан) означава се са 0, периоди који претходе са -1, -2 ..., а периоди који следе са +1, +2 ...

Код серија са парним бројем података, два средишња временска периода се означавају са -0,5 и +0,5, док се сви временски периоди који претходе -0,5 означавају са -1,5, -2,5..., а периоди који следе након +0,5 са +1,5, +2,5 ...

Из тога следи да је $\sum x_i = 0$.

Стандардна грешка тренда

Да би се израчунало просечно одступање оригиналних података од моделских вредности тренда, користи се *стандардна грешка тренда*. Стандардна грешка тренда може се израчунати на основу појединачних одступања оригиналних од оцењених вредности, према следећем обрасцу:

$$Sy_i = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n-k}}$$

где је k број оцењених параметара, односно код модела линеарног тренда k =2.

Пример 1.

На основу података о висини остварених прихода (у милионима динара) у периоду од 2008-2016. године:

а) Одредити функцију линеарног тренда и графички је приказати;

б) Израчунати стандардну грешку тренда.

Табела 1.

Година	Остварени приходи
2008	38,7
2009	38,2
2010	35,9
2011	32,1
2012	23,2
2013	18,9
2014	17,9
2015	16,9
2016	16,3

Решење:

а) Први корак у одређивању функције линеарног тренда јесте оцена непознатих параметара b_0 и b_1 на основу следећих образаца:

$$b_0 = \bar{y} = \frac{\sum y_i}{n}$$

$$b_1 = \frac{\sum x_i y_i}{\sum x_i^2}$$

Да бисмо израчунали непознате коефицијенте, неопходно је „кодирати“ године, односно уместо година ставити одговарајуће бројеве у зависности од тога да ли је посматрани скуп података паран или непаран. Пошто је посматран временски период од 2008. до 2016. године, односно укупно 9 година, закључујемо да је $n = 9$, односно да је посматрани скуп података непаран. Табелу попуњавамо тако што прво уместо средишње године, 2012., пишемо 0:

Табела 2.

Година	Остварени приходи (y_i)	x_i
2008	38,7	-4
2009	38,2	-3
2010	35,9	-2
2011	32,1	-1
2012	23,2	0
2013	18,9	+1
2014	17,9	+2
2015	16,9	+3
2016	16,3	+4
укупно	238,1	0

Затим израчунавамо вредности које су нам неопходне за оцену непознатих параметара b_0 и b_1.

Табела 3.

Година	Остварени приходи (y_i)	x_i	$x_i\ y_i$	x_i^2
2008	38,7	-4	-154,8	16
2009	38,2	-3	-114,6	9
2010	35,9	-2	-71,8	4

2011	32,1	-1	-32,1	1
2012	23,2	0	0	0
2013	18,9	1	18,9	1
2014	17,9	2	35,8	4
2015	16,9	3	50,7	9
2016	16,3	4	65,2	16
укупно	238,1	0	-202,7	60

Добијене вредности замењујемо у формулама:

$$b_0 = \bar{y} = \frac{\sum y_i}{n} = \frac{238{,}1}{9} = 26{,}46$$

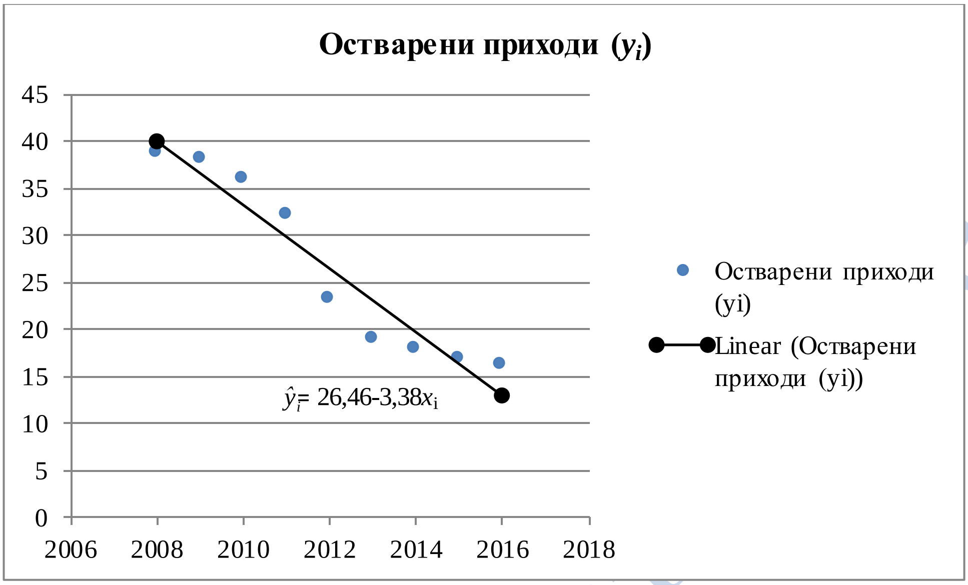
$$b_1 = \frac{\sum x_i y_i}{\sum x_i^2} = -\frac{202{,}7}{60} = -3{,}38$$

Оцењени модел линеарног тренда гласи:

$$\hat{y}_i = b_0 + b_1 \cdot x_i = 26{,}46 - 3{,}38 \cdot x_i$$

Оцењена вредност коефицијента b_0 указује да је просечна висина остварених прихода у периоду од 2008-2016. године износила 26.460.000 динара, док вредност коефицијента b_1 указује да су остварени приходи у посматраном периоду сваке године у просеку опадали за 3.380.000 динара (коефицијент b_1 је негативан, зато је у питању просечан пад).

Да би се функција линеарног тренда графички приказала, у координатни систем је потребно унети оригиналне податке, односно вредности оствареног прихода по годинама, а затим се на основу одабране две вредности (најчешће су то најмања и највећа кодирана вредност године) одређују координате на основу којих се уцртава линија тренда, односно:



Слика 1. Дијаграм линеарног тренда

Линија тренда је уцртана на основу следећих координата:

$$\hat{y}_i = b_0 + b_1 \cdot x_i = 26{,}46 - 3{,}38 \cdot (-4) = 39{,}98$$

Ова вредност се уписује наспрам 2008. године и представља почетну координату.

Крајња координата се уписује наспрам 2016. године и одређује се на следећи начин:

$$\hat{y}_i = b_0 + b_1 \cdot x_i = 26{,}46 - 3{,}38 \cdot 4 = 12{,}94$$

б) Да би се израчунала стандардна грешка линеарног тренда, потребно је израчунати појединачна одступања оригиналних података од оцењених вредности, односно:

Табела 4.

Година	Остварени приходи (y_i)	x_i	$\hat{y}_i$	$y_i - \hat{y}_i$	$(y_i - \hat{y}_i)^2$
2008	38,7	-4	39,98	-1,28	1,6384
2009	38,2	-3	36,6	1,6	2,56
2010	35,9	-2	33,22	2,68	7,1824
2011	32,1	-1	29,84	2,26	5,1076
2012	23,2	0	26,46	-3,26	10,6276

2013	18,9	1	23,08	-4,18	17,4724
2014	17,9	2	19,7	-1,8	3,24
2015	16,9	3	16,32	0,58	0,3364
2016	16,3	4	12,94	3,36	11,2896
укупно	238,1	0	238,14	-	59,4544

Израчунате вредности замењујемо у обрасцу:

$$Sy_i = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n-k}} = \sqrt{\frac{59{,}4544}{9-2}} = \sqrt{8{,}49} = 2{,}91$$

Просечно одступање оствареног прихода у односу на вредности предвиђене функцијом тренда износи 2.910.000 динара.

ЕКСПОНЕНЦИЈАЛНИ ТРЕНД

Уколико дугорочна тенденција појаве показује стабилну релативну промену, реч је о *експоненцијалном тренду*. Такође, сматра се да је присутан експоненцијални тренд када се подаци у временској серији повећавају тако да је процентуална разлика промене од вредности до вредности константна.

Оцењени модел

Теоријска функција експоненцијалног тренда гласи:

$$y_i = \beta_0 + \beta_1^{x_i}$$

а оцењена:

$$\hat{y}_i = b_0 \cdot b_1^{x_i}$$

где је:

$\hat{y}_i$ – оцена просечних вредности посматране појаве;

x_i – „кодиране“ године;

b_0 и b_1 – оцене параметара β_0 и β_1.

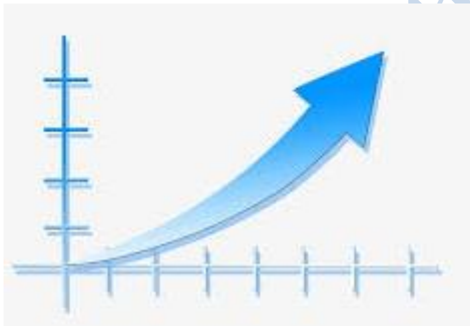
Да би се линеаризовао експоненцијалан однос између две променљиве врши се *логаритмовање*, па логаритмована функција експоненцијалног тренда гласи:

$$\log \hat{y}_i = \log b_0 + x_i \log b_1$$

Логаритмоване вредности непознатих коефицијената рачунају на следећи начин:

$$\log b_0 = \frac{\sum \log y_i}{n}$$

$$\log b_1 = \frac{\sum x_i \cdot \log y_i}{\sum x_i^2}$$



Експоненцијална стопа раста

Експоненцијална стопа раста показује просечан раст или пад посматране појаве у одговарајућем временском периоду. Може се израчунати на основу следећег обрасца:

$$r_e = (b_1 - 1) \cdot 100 = \%$$

Уколико је добијени резултат негативан, реч је о просечном паду, уколико је позитиван, посматрана појава остварује просечан раст.

Пример 2.

На основу података о потрошњи угља (у тонама) у једној фабрици у периоду од 2005-2011. године:

а) Одредити функцију експоненцијалног тренда;

б) Израчунати експоненцијалну стопу раста.

Табела 5.

Година	Потрошња угља
2005	475
2006	465
2007	451
2008	442
2009	431
2010	420
2011	410

Решење:

а) Први корак у одређивању функције линеарног тренда јесте оцена непознатих параметара b_0 и b_1 на основу следећих образаца:

$$\log b_0 = \frac{\sum \log y_i}{n}$$

$$\log b_1 = \frac{\sum x_i \cdot \log y_i}{\sum x_i^2}$$

Да бисмо израчунали непознате коефицијенте, неопходно је „кодирати“ године, односно уместо година ставити одговарајуће бројеве у зависности од тога да ли је посматрани скуп података паран или непаран. Пошто је посматран временски период од 2005. до 2011. године, односно укупно 7 година, закључујемо да је n = 7, односно да је посматрани скуп података непаран. Табелу попуњавамо тако што прво уместо средишње године, 2008., пишемо 0:

Табела 6.

Година	Потрошња угља (y_i)	x_i
2005	475	-3
2006	465	-2
2007	451	-1
2008	442	0
2009	431	+1
2010	420	+2
2011	410	+3
укупно	3.094	0

Затим израчунавамо вредности које су нам неопходне за оцену непознатих параметара b_0 и b_1.

Табела 7.

Година	Потрошња угља (y_i)	x_i	x_i^2	$\log y_i$	$x_i \cdot \log y_i$
2005	475	-3	9	2,676694	-8,03008
2006	465	-2	4	2,667453	-5,33491
2007	451	-1	1	2,654177	-2,65418
2008	442	0	0	2,645422	0

2009	431	1	1	2,634477	2,634477
2010	420	2	4	2,623249	5,246499
2011	410	3	9	2,612784	7,838352
укупно	3,094	0	28	18,51426	-0,29984

Добијене вредности замењујемо у формулама:

$$log\, b_0 = \frac{\sum log\, y_i}{n} = \frac{18{,}51426}{7} = 2{,}6449 => b_0 = 441{,}46$$

$$log\, b_1 = \frac{\sum x_i \cdot log\, y_i}{\sum x_i^2} = -\frac{0{,}29984}{28} = -0{,}0107 => b_1 = 0{,}976$$

Напомена: Да би се добиле вредности непознатих коефицијената неопходно је извршити *антилогаритмовање* добијених вредности 2,6449 и -0,0107.

Оцењени модел експоненцијалног тренда гласи:

$$\hat{y}_i = b_0 \cdot b_1^{x_i} = 441{,}46 \cdot 0{,}976^{x_i}$$

б) Експоненцијалну стопу раста можемо израчунати на основу већ израчунатог коефицијента b_1:

$$r_e = (b_1 - 1) \cdot 100 = (0{,}976 - 1) \cdot 100 = -2{,}4\%$$

Потрошња угља у посматраном периоду просечно је опадала за 2,4%.

Др Наташа Папић-Благојевић, проф.