

XIII.1. Pojam vremenske serije

Vremensku seriju u najopštijem slučaju predstavlja niz vremenski uređenih rezultata posmatranja. U ovom tekstu pod vremenskom serijom podrazumevaćemo uređeni skup rezultata posmatranja nekog procesa, pri čemu je skup uređen hronološki, tj. dobijen posmatranjem jednog obeležja na istim jedinicama posmatranja u jednakim vremenskim intervalima ili u ekvidistantnim vremenskim tačkama⁵⁵.

Vremenske serije mogu biti:

1. **Neprekidne** - vrednosti vremenske serije određene su u svakom proizvoljnom vremenskom trenutku (npr. zapis elektroencefalograma) ili
2. **Diskretne** - vrednosti vremenske serije određene su u zadatim prekidnim vremenskim trenucima (na primer: broj pacijenata koji dođu psihologu na svakih sat vremena, broj udesa u nekom gradu po danima, broj reprodukovanih prethodno naučenih slogova na svaki sat po završenom učenju). Diskretne vremenske serije se, pak, mogu podeliti na momentne i intervalne. Kod **momentnih** vremenskih serija vrednosti vremenske serije posmatraju se u zadatim, ekvidistantnim vremenskim tačkama. U takve serije spada, na primer, broj tačno reprodukovanih reči posle svakog sata od trenutka kada su sve reči naučene. **Intervalne** vremenske serije su one u kojima se vrednosti vremenske serije dobijaju usrednjavanjem pokazatelja za određeni vremenski interval. Primer takvih vremenskih serija bila bi prosečna količina potrošenih lekova za svaki mesec tokom određenog vremenskog perioda.

Teorijski posmatrano vrednost vremenske serije u trenutku t , u oznaci v_t , sastoji se iz dve komponente: determinisane ili sistematske (v_t^*) i slučajne komponente ili šuma (γ_t):

$$v_t = v_t^* + \gamma_t$$

Statističke metode analize vremenskih serija mogu se podeliti u dve grupe:

⁵⁵ Posebna oblast analize vremenskih serija je ispitivanje povezanosti dveju ili više varijabli tako da se vrednosti tih varijabli snimaju na istim jedinicama posmatranja tokom vremena. Tako, na primer, moguće je ispitivati efekat stepena nezaposlenosti na nivo kriminala snimanjem promena obe ove pojave tokom dužeg niza godina. Budući da između sukcesivnih vrednosti na tim varijablama u ovakvom slučaju postoji tzv. autokorelisanost reziduala, tj. grešaka, nije dozvoljeno za analizu efekata jedne varijable na drugu primenjivati standardni regresioni model.

1. Metode koje analizi vremenskih serija prilaze iz domena frekvenci - harmonijska i spektralna analiza;
 2. Metode koje analizi vremenskih serija prilaze iz domena vremena. U ove metode spadaju modeli razvojnih tendencija, metode za analizu međusobnih odnosa vremenskih serija i metode za predviđanje vremenskih serija u funkciji vremena.
- U ovom tekstu prikazani su osnovni postupci koji analizi vremenskih serija prilaze iz domena vremena. Od mnoštva pristupa i modela u okviru ove grupe postupaka prikazaćemo samo metode dekompozicije i osnovne postupke izravnavanja (eng. smoothing) vremenske serije.

Osnovni ciljevi u analizi vremenskih serija su otkrivanje mehanizma koji generiše vremensku seriju, tj. definisanje adekvatnog modela za opis vremenske serije i prognoza budućih vrednosti serije.

XIII. 2. Osnovni pojmovi neophodni za razumevanje analize vremenskih serija

Operator docnje (eng. lag operator)

Ako je Y_t vrednost vremenske serije u trenutku t , a Y_{t-k} vrednost vremenske serije u trenutku koji prethodi vremenu t za k perioda ili momenata, tada se operator docnje, u oznaci B^k , definiše na sledeći način:

$$B^k Y_t = Y_{t-k}, \quad \text{pri čemu je } B^0 Y_t = Y_t$$

Ovaj operator, dakle, vrednost vremenske serije u trenutku t pomera za k perioda unazad.

Operator k-te difference (eng. difference operator)

Operator k-te difference, u oznaci D_k , definiše se na sledeći način:

$$D_k = 1 - B^k$$

pri čemu je B^k operator docnje. Očigledno je da je

$$D_k = Y_t - Y_{t-k}$$

jer je $(1 - B^k)Y_t = Y_t - B^k Y_t = Y_t - Y_{t-k}$.

U analizi vremenskih serija najčešće se koristi operator prve difference, tj. operator D_1 . Najčešće se jedinica u oznaci ovog operatora izostavlja. Dakle, operator prve difference, u oznaci D , definiše se na sledeći način:

$$D = 1 - B = Y_t - Y_{t-1}$$

Periodičnost (eng. periodicity)

Periodičnost vremenske serije odnosi se na ciklična variranja vremenskih jedinica unutar nekog vremenskog perioda. Na primer, kvartali imaju periodičnost jednaku 4, a meseci periodičnost jednaku 12 unutar godine. Dani imaju periodičnost koja iznosi 7 unutar nedelje, a minuti imaju periodičnost jednaku 60 unutar časa. Isto tako, periodičnost sati unutar dana iznosi 24. Međutim, radni sati imaju periodičnost 8 unutar dana, ako je radno vreme 8 sati.

Operator sezonske difference (eng. seasonal difference operator)

Operator sezonske difference, u oznaci D_s , definiše se na sledeći način:

$$D_s = 1 - B^s$$

pri čemu je s periodičnost vremenske serije, tj. broj sezona. Na primer, za mesečne vremenske serije s je 12, a za kvartalne s je 4. Prema tome, za mesečne vremenske serije $D_{12} Y_t = (1 - B^{12})Y_t = Y_t - Y_{t-12}$.

Diferentovanje (eng. differencing)

Postupak primene operatora prve difference na podatke vremenske serije zovemo diferentovanje.⁵⁶

Stepen ili red diferentovanja

Stepen diferentovanja govori o tome koliko se puta vremenska serija transformiše tako što se od vrednosti serije u trenutku t oduzima vrednost serije u prethodnom trenutku, tj. koliko se puta primenjuje operator prve difference. Tako, na primer ako vremensku seriju čini sledeći niz podataka snimljenih u vremenskim trenucima od 1 do 10: 12, 15, 20, 25, 40, 60, 90, 120, 160, 220 prvom primenom operatora prve difference dobijamo niz 3, 5, 5, 15, 20, 30, 30, 40, 60. Sledeća primena operatora prve difference daje sledeći niz 2, 0, 10, 5, 10, 0, 10, 20. Budući da smo operator prve difference primenili dva puta stepen ili red diferentovanja je dva. Najčešće se u analizi vremenskih serija koristi diferentovanje prvog i drugog stepena ili reda.

Sezonsko diferentovanje

⁵⁶ Bez obzira na to što se u statističkoj literaturi na našem jeziku ovaj postupak najčešće zove diferenciranje mi smo se odlučili za ovaj jezički neubičajen oblik kako bismo ovaj postupak razlikovali od diferenciranja u diferencijalnom računu.

Diferentovanje sezonskih vremenskih serija vrši se korišćenjem operatora sezonske difference. S obzirom na definiciju ovog operatora, diferentovanje sezonskih serija se izvodi tako što se od vrednosti serije u datom trenutku oduzima njoj odgovarajuća vrednost s obzirom na periodičnost serije, tj. broj sezona, a ne vrednost koja joj neposredno prethodi. Tako, na primer u seriji koja sadrži podatke za sve mesece iz 1991. i 1992. godine i, prema tome, mesečnu periodičnost diferentovanje prvog reda se vrši tako što se od vrednosti za januar 1992. godine oduzima vrednost iz januara 1991. godine (a ne vrednost iz decembra 1991 godine), i tako redom.

Izravnavanje (eng. smoothing)

Izravnavanje vremenske serije je postupak kojim se iz vremenske serije odstranjuju slučajne, nesistematske varijacije kako bi se lakše identifikovao trend, tj. opšta tendencija promene vremenske serije u funkciji vremena.

Pokretna sredina (eng. moving average)

Pokretna sredina ili pokretni prosek predstavlja jednako ili različito ponderisan prosek određenog podatka vremenske serije i određenog broja podataka u njegovoj okolini. Broj podataka vremenske serije koji ulaze u računanje pokretnog proseka naziva se rasponom (eng. span). Tako, ako je raspon jednak tri tada se pokretna sredina kojom će se pri izravnavanju vremenske serije zameniti podatak Y_t , računa kao prosek podataka Y_{t-1} , Y_t i Y_{t+1} .

Dužina horizonta prognoziranja

Dužina horizonta prognoziranja, u oznaci $Y_n^*(h)$, govori o broju budućih vremenskih trenutaka za koje, na osnovu raspoloživih podataka za n vremenskih trenutaka, vršimo prognozu vremenske serije. Ukoliko, na primer, imamo mesečne podatke za sve mesece iz 1991. i 1992. godine, a h je 4, tada ćemo dužinu horizonta prognoziranja označiti sa $Y_{24}^*(4)$. To znači da na osnovu podataka za 24 meseca iz 1991. i 1992. godine, treba da napravimo prognozu za januar, februar, mart i april 1993. godine.

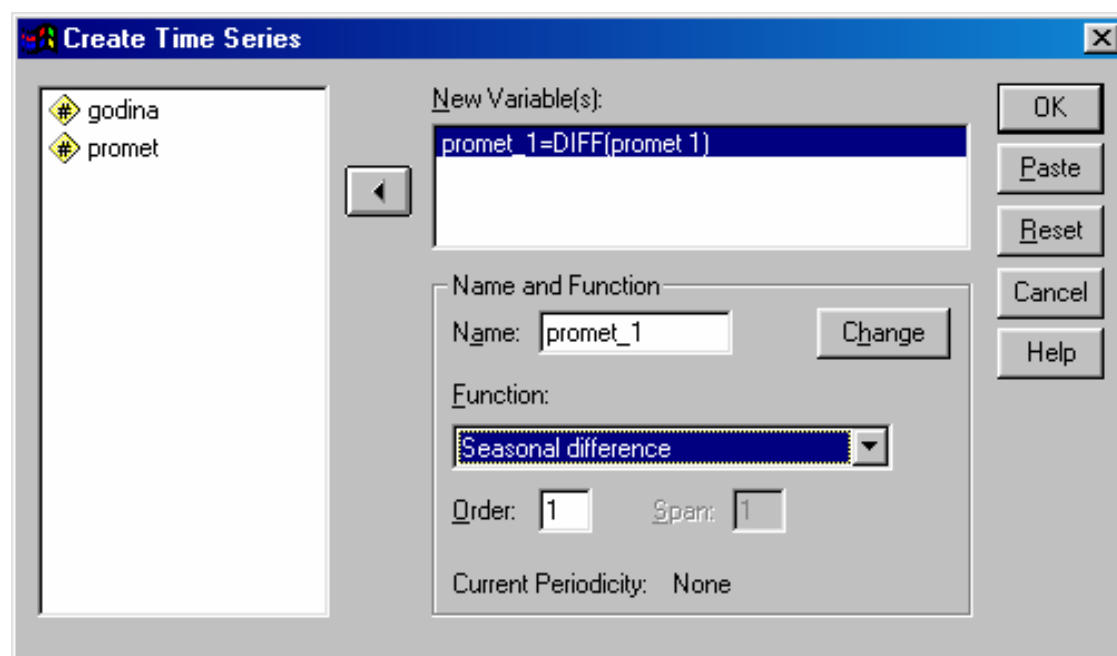
Podaci za analizu vremenskih serija u programu SPSS moraju biti organizovani tako da su u redovima opservacije, tj. rezultati serije za jedan vremenski interval ili vremenski trenutak. Da bi se definisala periodičnost za vremenske serije u koje je uključena sezonska komponenta u meniju **Data/Define Dates** mora se definisati, pored varijable u kojoj su vrednosti serije za n vremenskih tačaka i **datumska** varijabla (**Date variable**) koja sadrži podatak o tome kakva je periodičnost vremenske serije. U dijalogu **Define Dates** treba definisati periodičnost klikom na odgovarajuću opciju u spisku **Cases Are**, a zatim u okviru **First Case is**

početni trenutak, tj. vremenski trenutak od kojeg počinje serija. Na primer, ako je periodičnost serije jedan kvartal i ako serija počinje sa prvim kvartalom 1990. godine, u spisku **Cases Are** treba kliknuti na **Years,quarters**, a u okviru **First Case is** u polje **Year** upisati 1990 i u polje **Quarter** jedinicu. Posle završnog klika na dugme **OK** program će sam definisati odgovarajuće varijable u tabeli podataka.

Ukoliko se želi eliminisati datumska varijabla koju je program napravio u tabeli za podatke treba u spisku **Cases are** odabrati opciju **Not dated** i potom kliknuti na dugme **OK**.

Diferentovanje vremenske serije i sezonsko diferentovanje u programu SPSS izvodi se u meniju **Transform/Create Time Series** (slika 12). U polje **New Variable(s)**

*Slika 12: Prozor za dijalog procedure **Create Time Series***



treba prebaciti ime varijable u kojoj se nalaze podaci za vremensku seriju. Zatim u okviru **Name and Function** u polje **Name** treba upisati ime varijable u kojoj će biti upisane transformisane vrednosti vremenske serije posle diferentovanja. Klik na strelicu sa desne strane polja **Function** omogućava pojavljivanje spiska funkcija kojima se može transformisati vremenska serija. Za diferentovanje ili sezonsko diferentovanje treba odabrati opciju **Difference**, odnosno **Seasonal difference**. U polju **Order** definiše se red ili stepen diferentovanja upisivanjem cifre koja

odgovara stepenu diferentovanja. Ukoliko se u poljima u okviru **Name and Function** vrše bilo kakve izmene potrebno je posle ovih izmena kliknuti na dugme **Change**. Da bi se moglo definisati sezonsko diferentovanje neke vremenske serije pored varijable u kojoj su vrednosti vremenske serije u fajlu sa podacima mora postojati i datumska varijabla koja se prethodno kreira u meniju **Data/Define Dates**. U primeru na slici 12 definisano je diferentovanje prvog reda vremenske serije iz varijable **promet**. Vrednosti diferentovane serije u ovom primeru biće upisane u novoj varijabli **promet_1**.

XIII.4. Grafički prikaz jedne vremenske serije

Za grafičko prikazivanje jedne vremenske serije mogu se koristiti svi grafički postupci za kvantitativne varijable koji su prikazani u odeljku **VI.2**: histogram, stablogram i kutijasti dijagram. Ipak, najčešće se jedna vremenska serija prikazuje dijagramom niza (eng. sequence plot), tj. vremenskim dijagramom (eng. time plot). Na apscisi vremenskog dijagrama su vremenski trenuci ili periodi u kojima je snimana vremenska serija a na ordinati ovog dijagrama su vrednosti vremenske serije. Ovim dijagramom mogu se prikazati izvorne netransformisane vrednosti vremenske serije ali i transformisane vrednosti koje su dobijene posle diferentovanja ili izravnavanja vremenske serije. Vremenski dijagram može biti veoma koristan u uočavanju oblika trenda vremenske serije kao i postojanja cikličnih i sezonskih kolebanja.

Vremenski dijagram se u programu SPSS pravi u meniju **Graphs/Sequence**. U polje **Variables** treba prebaciti ime varijable koja sadrži vremensku seriju, a u polje **Time Axis Label** ime varijable u kojoj su oznake vremenskih tačaka u kojima su snimane vrednosti vremenske serije, ako takva varijabla postoji. Popunjavanje polja **Time Axis Label** nije obavezno.

XIII.4. Analiza trenda

U otkrivanju trenda, tj. opšte tendencije promene vremenske serije u funkciji vremena koriste se različiti "empirijski postupci" iza kojih ne stoji jedan opšti teorijski model kao i strukturni modeli. U ovom tekstu prikazaćemo samo osnovne postupke izravnavanja vremenske serije u okviru empirijskog pristupa.

Izravnavanje vremenske serije

Postupci izravnavanja vremenske serije polaze od pretpostavke da u seriji pored slučajnih kolebanja, tj. varijacija postoji opšta zakonitost, trend serije. Cilj izravnavanja je da se slučajna variranja priguše kako bi na taj način postala uočljivija osnovna tendencija vremenske serije. Izravnata serija se potom dalje koristi u prognozama budućih vrednosti vremenske serije.

Izravnavanje pokretnim sredinama

Ovaj postupak izravnavanja najčešće se koristi kako bi se uočio trend kod pojava sa sezonskim i cikličnim kolebanjima. Postupak se sastoji u tome da se originalni podaci vremenske serije zamene pokretnim prosecima, tj. aritmetičkim sredinama date vrednosti serije i određenog broja vrednosti koje su ispred i iza nje u nizu.

Najveći nedostatak ovog postupka izravnavanja je u tome što skraćuje vremensku seriju za prvih nekoliko i poslednjih nekoliko tačaka, zavisno od raspona pokretnog proseka. To je poseban problem u situacijama kada se želi prognoziranje budućeg kretanja serije.

Izravnavanje vremenske serije pokretnim prosecima izvodi se u programu SPSS korišćenjem menija **Transform/Create Time Series** (slika 12). U polje **New Variable(s)** treba prebaciti ime varijable u kojoj se nalaze početni neizravnati podaci za vremensku seriju. Zatim u okviru **Name and Function** u polje **Name** treba upisati ime varijable u kojoj će biti upisane transformisane vrednosti vremenske serije posle izravnavanja. Klik na strelicu sa desne strane polja **Function** omogućava pojavljivanje spiska funkcija kojima se može transformisati vremenska serija. Za izravnavanje serije pokretnim prosecima treba odabrati opciju **Centered moving average**, a u polju **Span** treba definisati raspon, tj. broj podataka početne vremenske serije koji ulaze u računanje pokretnog proseka. Tako, na primer, ako želimo da se dati originalni podatak vremenske serije zameni aritmetičkom sredinom tog podatka, podatka koji mu prethodi i podatka koji sledi odmah iza njega u polju **Span** treba upisati cifru **3**. Posle neophodnih promena u poljima u okviru **Name and Function** potrebno je kliknuti na dugme **Change**. varijable **promet**. Vrednosti izravunate serije posle klika na dugme **OK** biće upisane u radnom fajlu za podatke u varijabli čije je ime definisano u polju **Name** okvira **Name and Function**.

Eksponecijalno izravnavanje

Osnovna karakteristika postupaka ekspanencijalnog izravnavanja (eng. exponential smoothing) je da se zasniva na ponderisanim pokretnim prosecima tako da veličina pondera ekspanencijalno opada sa starošću podatka u seriji. Dakle, veći ponder, tj. važnost pri računanju pokretnih proseka daje se novijim podacima. Postoje različiti

modeli eksponencijalnog izravnavanja (Holtov, Wintersov...) koji se međusobno razlikuju prema pretpostavkama o trendu i sezonskoj komponenti serije.

Postupci eksponencijalnog izravnavanja izvode se u programu SPSS korišćenjem menija **Analyze/Time Series/Exponential Smoothing**. Osnovni problemi koje pri tome treba rešiti jesu izbor najpodesnijeg modela za seriju, tretiranje sezonskih varijacija i izbor parametara izravnavanja.

Aproksimacija trenda matematičkim modelom

Problem aproksimacije trenda polinomom može se formalno definisati na sledeći način:

$$\mathbf{Tb} = \mathbf{y} - \mathbf{g} \mid \mathbf{g}^t \mathbf{g} = \min$$

$$\mathbf{b} = ?$$

Matrica \mathbf{T} je matrica reda $f \times f$ čija prva kolona ima f jedinica, a ostalih $f - 1$ kolona služe za određivanje koeficijenata polinoma. U preostalih $f - 1$ vektora su: u prvom sledećem vektoru celi brojevi od 1 do f , zatim kvadrati ovih brojeva, treći stepeni celih brojeva i tako redom. Izraz $\mathbf{g}^t \mathbf{g}$ označava sumu kvadriranih reziduala (grešaka), tj. zbir kvadriranih razlika empirijski dobijenih vrednosti vremenske serije i očekivanih vrednosti serije na osnovu modela.

Problem je, dakle, naći osnovnu tendenciju vremenskog niza, složaj vremenske serije, zakonitost po kojoj se menjaju rezultati vremenske serije u funkciji vremena, tj. naći što jednostavniju matematičku funkciju koja što je moguće bolje opisuje zakonitost promene vrednosti vremenske serije u funkciji vremena. Potrebno je opisati što je moguće tačnije skup tačaka koje definišu dobijene vrednosti vremenske serije u vremenskim tačkama $t_1, t_2, t_3, \dots, t_f$ polinomom što je moguće nižeg stepena, tj. što je moguće jednostavnom matematičkom funkcijom. Opšti oblik polinoma za aproksimaciju trenda je:

$$y = \sum_{p=0}^s b_p t^p + g = b_0 + b_1 t + b_2 t^2 + \dots + b_s t^s + g$$

Bez obzira na to što se raspored od f tačaka može savršeno opisati polinomom stepena $f-1$, ne treba zaboraviti da uvođenje složenijeg modela podrazumeva dodavanje novog parametra i (zbog neophodnosti njegove ocene na osnovu podataka) gubitka jednog stepena slobode.

Rešenje problema aproksimacije trenda vremenske serije po principu najmanjih kvadrata (za polinome reda manjeg od 5) dobija se na sledeći način:

$$\mathbf{b} = (\mathbf{T}^t \mathbf{T})^{-1} \mathbf{T}^t \mathbf{y}$$

pri čemu je \mathbf{b} vektor koeficijenata polinoma.

Očekivane vrednosti varijable y u vremenskim tačkama $t_1, t_2, t_3, \dots, t_f$ nalaze se u vektoru $\mathbf{y}^* = \mathbf{Tb} = \mathbf{T}(\mathbf{T}^t \mathbf{T})^{-1} \mathbf{T}^t \mathbf{y}$. Deljenjem $\mathbf{g}^t \mathbf{g}$, tj. zbira kvadriranih reziduala sa vrednošću izraza $f - p - 1$, pri čemu je p je broj parametara, tj. koeficijenata koje ocenjujemo na osnovu podataka, dobijamo varijansu reziduala, u oznaci S_g^2 :

$$S_g^2 = \frac{\sum_{t=1}^f g_t^2}{f - p - 1} = \frac{\sum_{t=1}^f (y_t - y_t^*)^2}{f - p - 1}$$

Kao mere tačnosti, tj. podesnosti (engl. fit) modela u aproksimiranju rasporeda empirijskih tačaka koriste se koeficijent korelacije predviđenih i empirijski dobijenih vrednosti vremenske serije i standardna greška ocene, tj. standardna devijacija distribucije reziduala ili kvadratni koren varijanse reziduala. Za određivanje vektora u \mathbf{T} koji značajno doprinose objašnjenju trenda može poslužiti parcijalna korelacija pojedinih elemenata (komponenti) polinoma i empirijski dobijenih vrednosti vremenske serije.

Predviđanje budućih vrednosti vremenske serije

Za razliku od interpolacije koja se odnosi na određivanje (na osnovu modela) vrednosti vremenske serije u intervalu između sukcesivnih vrednosti u ispitivanom rasponu vremenskih tačaka, ekstrapolacija se odnosi na prognozu vrednosti zavisne varijable u nekom budućem vremenskom trenutku. Pri ekstrapolaciji treba biti veoma oprezan jer je ona zavisna od toga da li model važi i za buduća kretanja vrednosti vremenske serije.

Odgovor na pitanje kojim polinomom najnižeg stepena ili kojom najjednostavnijom matematičkom funkcijom uopšte možemo dobro da aproksimiramo trend neke vremenske serije u programu SPSS možemo potražiti odabirom menija **Analyze/Regression/Curve Estimation**. U dijalogu **Curve Estimation** kao **Dependent(s)** varijablu treba uneti ime zavisne, kriterijumske varijable u kojoj se nalaze podaci vremenske serije, a u okviru **Independent** odabrati vreme kao prediktor (opcija **Time**). Ukoliko želimo grafički prikaz rezultata analize možemo uključiti opciju **Plot Models**. U okviru **Models** treba odabrati željeni model i uključiti opciju **Display Anova Table** ispod ovog okvira. Ponuđeni modeli (polinomski i drugi) u okviru **Models** su:

Linearni (**Linear**): $Y_t = b_0 + b_1 t$

Kvadratni (**Quadratic**): $Y_t^* = b_0 + b_1t + b_2t^2$

Kubični (**Cubic**): $Y_t^* = b_0 + b_1t + b_2t^2 + b_3t^3$

Eksponecijalni (**Exponential**): $Y_t^* = b_0e^{b_1t}$

Modifikovani eksponencijalni (**Compound**): $Y_t^* = b_0b_1^t$

Model rasta (**Growth**): $Y_t^* = e^{b_0 + b_1t}$

Logaritamski (**Logarithmic**): $Y_t^* = b_0 + b_1 \ln(t)$

S model (**S**): $Y_t^* = e^{b_0 + b_1/t}$

Hiperbolički (**Inverse**): $Y_t^* = b_0 + b_1/t$

Stepeni ili parabolički (**Power**): $Y_t^* = b_0 t^{b_1}$

Logistički (**Logistic**): $Y_t^* = 1 / (1/u + b_0 b_1^t)$. U ovom modelu oznakom u obeležena je gornja granična vrednost koja mora biti pozitivna i veća od najveće vrednosti u kriterijumskoj varijabli.

Predviđena vrednost vremenske serije označena je u gornjim jednačinama sa Y_t^* , prediktorska varijabla, tj vreme je t , dok su b_0 , b_1 , b_2 , i b_3 ocene parametara, tj. koeficijenti modela.⁵⁷

U opciji **Save** može se zahtevati zapisivanje u aktivnom fajlu sa podacima novih varijabli sa predviđenim (na osnovu modela) vrednostima vremenske serije (**Predicted values**), reziduala (**Residuals**) i 90%, 95% ili 99% interval poverenja za predviđene vrednosti (odabirom opcije **Prediction intervals** i odabirom odgovarajućeg broja u polju ispred **% Confidence interval**). Gornje i donje granice ovog intervala u aktivnom fajlu će biti sačuvane kao posebne varijable.

U delu ispisa koji nosi naslov **Variables in the Equation** date su u koloni **B** ocene parametara (koeficijenata) matematičkog modela trenda.

Adekvatnost matematičke funkcije (modela) za opis trenda vremenske serije može se odrediti poređenjem vrednosti statistika **Multiple R** (korelacija predviđenih modelom i empirijskih vrednosti vremenske serije) i **Standard Error** (standardna devijacija distribucije reziduala) za različite modele kao i pregledanjem grafika, tj. na osnovu ispitivanja koliko je kriva, definisana modelom, blizu empirijskim tačkama. Što je korelacija veća a standardna greška manja to je model adekvatniji. Opređeljujemo se za matematički složeniji model samo ako je znatno superiorniji od jednostavnijeg u opisu trenda vremenske serije.

Primer:

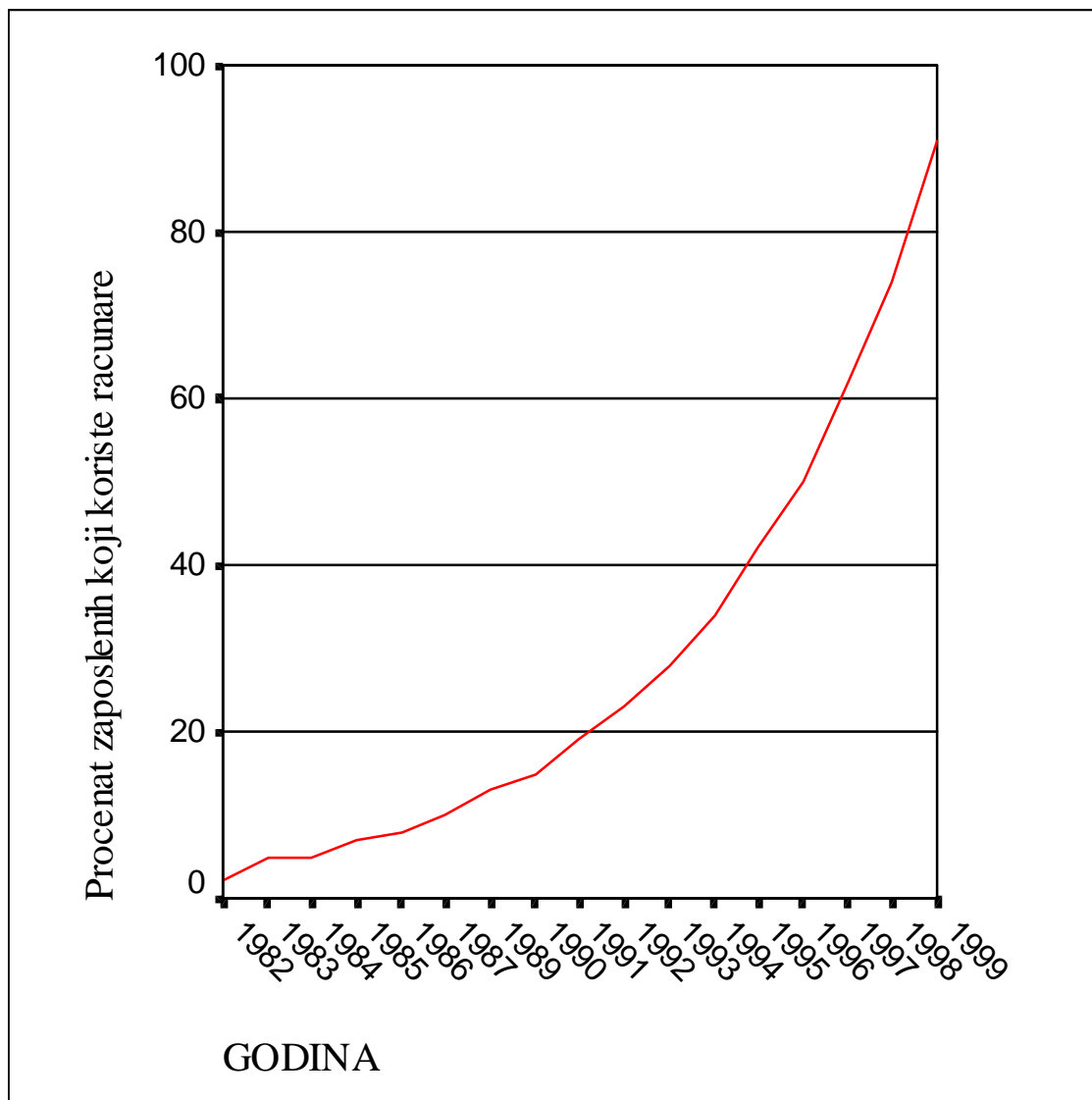
Procenti zaposlenih koji redovno koriste računar u svom poslu u jednoj firmi po kalendarskim godinama u periodu od 1982 do 1999 prikazani su na vremenskom dijagramu dobijenom u proceduri **Graphs/Sequence** programa SPSS (Grafik 13.1). Potrebno je, na osnovu podataka koji su poslužili za crtanje grafika odabrati

⁵⁷ U svim modelima e je oznaka osnove prirodnih logaritama, koja je, kao što je poznato, približno jednaka 2.718. Inače, e se matematički definiše na sledeći način:

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n$$

jednostavan matematički model koji dobro aproksimira trend ove pojave u ispitivanom vremenskom razdoblju.

Grafik 13.1: Procenat zaposlenih koji redovno koriste računar u svom poslu u jednoj firmi po kalendarskim godinama u periodu od 1982 do 1999



Na osnovu izvornih podataka koji su poslužili za crtanje grafika, napravljena je analiza kojom je trebalo aproksimirati trend pogodnim matematičkim modelom. Ispis iz procedure **Analyze/Regression/Curve Estimation** sa odabranim

potencijalnim modelima **Linear**, **Quadratic** i **Exponential** u okviru **Models**, i uključenom opcijom **Display Anova Table** izgleda ovako:

Curve Fit

MODEL: MOD_1.

—

Dependent variable.. PRKORRAC Method.. LINEAR

Listwise Deletion of Missing Data

Multiple R **,92857**
R Square ,86224
Adjusted R Square ,85306
Standard Error **10,20851**

Analysis of Variance:

	DF	Sum of Squares	Mean Square
Regression	1	9784,3235	9784,3235
Residuals	15	1563,2059	104,2137

F = 93,88709 Signif F = ,0000

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
Time	4,897059	,505397	,928570	9,690	,0000
(Constant)	-15,367647	5,178775		-2,967	,0096

—

Dependent variable.. PRKORRAC Method.. QUADRATI

Listwise Deletion of Missing Data

Multiple R **,99474**
R Square ,98952
Adjusted R Square ,98802
Standard Error **2,91511**

Analysis of Variance:

DF	Sum of Squares	Mean Square
----	----------------	-------------

Regression	2	11228,559	5614,2797
Residuals	14	118,970	8,4979

F = 660,66964 Signif F = ,0000

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
Time	-2,872291	,613190	-,544638	-4,684	,0004
Time**2	,431631	,033109	1,515789	13,037	,0000
(Constant)	9,235294	2,397614		3,852	,0018

—

Dependent variable.. PRKORRAC Method.. EXPONENT

Listwise Deletion of Missing Data

Multiple R	,99095
R Square	,98198
Adjusted R Square	,98078
Standard Error	,14997

Analysis of Variance:

	DF	Sum of Squares	Mean Square
Regression	1	18,384257	18,384257
Residuals	15	,337344	,022490

F = 817,45587 Signif F = ,0000

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
Time	,212272	,007424	,990950	28,591	,0000
(Constant)	2,649698	,201582		13,145	,0000

Uočavamo da eksponencijalni model najbolje opisuje trend ove pojave u ispitivanom vremenskom razdoblju. Premda je korelacija vrednosti predviđenih modelom i empirijskih vrednosti vremenske serije (**Multiple R**) podjednako visoka za kvadratni i eksponencijalni model, standardna greška (**Standard Error**) za eksponencijalni trend je znatno niža (0.15) od standardne greške za kvadratni model. Linearni model je prema oba pokazatelja neadekvatan. Prema tome, eksponencijalni model $Y_t^* = 2.65e^{0.21t}$ mogao bi biti model izbora za trend kretanja ispitivane pojave u datom vremenskom razdoblju. Treba uočiti da se za svaki model, u delu ispisa pod naslovom ----- **Variables in the Equation** -----, u koloni **B**, nalaze ocene parametara datog modela. Za eksponencijalni model $Y_t^* = b_0 e^{b_1 t}$ ocena parametra b_0 nalazi se u koloni **B** u redu **Constant**, a ocena parametra b_1 u redu **Time**.

XIII.5. Dekompozicija vremenske serije i isključenje pojedinih komponenti serije

Dekompozicija vremenske serije

Pri empirijskoj analizi promena određene pojave u funkciji vremena, varijacije vremenske serije možemo razložiti na četiri ključne komponente:

1. Opšta razvojna tendencija pojave u funkciji vremena – trend (t_t);
2. Ciklična kolebanja koja se ponavljaju u (često nejednakim) razmacima (c_t);
3. Sezonske varijacije koje se javljaju u razmacima manjim od jedne godine i ponavljaju se u dužem nizu godina (s_t)
4. Rezidualne (slučajne) varijacije (g_t)

Prve tri komponente čine sistematsku ili determinisanu komponentu vremenske serije. Postoje različite teorije o međusobnim odnosima pojedinih komponenti vremenskih serija:

I. Multiplikativna: $Y_t = t_t * c_t * s_t * g_t$;

II. Aditivna: $Y_t = t_t + c_t + s_t + g_t$;

III. Hibridna: $Y_t = t_t + (c_t * s_t * g_t)$ itd.

Isključenje trend komponente

Ponekada je potrebno iz vremenske serije isključiti trend komponentu. To je veoma često neophodno kada se računa korelacija između varijabli čiji podaci predstavljaju vremenske serije sa suprotnim trendovima. U tom slučaju bi računanje korelacije između varijabli iz kojih nisu isključeni trendovi dalo lažnu negativnu korelaciju mada među varijablama po prirodi postoji pozitivna povezanost. Jedan od klasičnih postupaka isključenja trend komponente polazi od određenog modela dekompozicije vremenske serije. Ukoliko je model o odnosima komponenti vremenske serije aditivni model, tada se ocenjene vrednosti trenda oduzimaju od vrednosti vremenske serije, a ukoliko je reč o multiplikativnom modelu tada se vrednost vremenske serije podeli ocenjenim vrednostima trenda.

U programu SPSS ocenjene vrednosti trenda mogu se, zavisno od postupka ocenjivanja (pokretni proseci u proceduri **Create Time Series** ili aproksimacija trenda matematičkom funkcijom u proceduri **Statistics/Regression/Curve**

Estimation) sačuvati kao varijabla i zatim komandom **Transform/Compute** napraviti isključivanje (oduzimanjem ili deljenjem) ocenjenih vrednosti trenda iz originalne vremenske serije.

Isključenje sezonske komponente – sezonsko izravnavanje ili desezoniranje vremenske serije

Za desezoniranje vremenske serije u upotrebi su dva pristupa: empirijski i modelski. Empirijski pristup zasniva se na ocenjivanju komponente trend-ciklus⁵⁸ i sezonske komponente pomoću pokretnih proseka, dok se u modelskom pristupu sezonska i nesezonska komponenta modeliraju korišćenjem strukturnih ili tzv. ARIMA modela. U ovom tekstu bavićemo se samo empirijskim pristupom⁵⁹. U empirijskom pristupu desezoniranju vremenska serija se dekomponuje na tri komponente: trend-ciklus, sezonsku i slučajnu (rezidualnu, iregularnu) komponentu. Sam postupak desezoniranja u ovom pristupu teče u nekoliko koraka:

- ✓ Komponenta trend-ciklus ocenjuje se izravnavanjem serije pomoću pokretnih proseka;
- ✓ Originalna vremenska serija se deli izravnatom serijom (ako se radi o multiplikativnom modelu) ili se od originalne serije oduzima izravnata serija (ako se radi o aditivnom modelu). Na taj način se iz originalne serije isključuje trend-ciklus komponenta i dobija "sezonsko-slučajna" komponenta;
- ✓ Sezonska komponenta se izoluje iz "sezonsko-slučajne" komponente računanjem tzv. "medijalnih" proseka za svaki vremenski period.

Desezoniranje se u programu SPSS izvodi pomoću menija **Analyze/Time Series/Seasonal Decomposition**. Naravno, desezoniranje će biti moguće samo ukoliko je u radnom fajlu sa podacima pored varijable u kojoj su podaci za vremensku seriju definisana i datumska varijabla postupkom koji smo objasnili u odeljku **XIII.2**. U okvir **Variable(s)** prozora za dijalog **Seasonal Decomposition** ubacuje se ime varijable koja sadrži podatke za vremensku seriju. U okviru **Model** ovog dijaloga treba odabrati odgovarajući model za odnos trend-ciklus i sezonske komponente (aditivni ili multiplikativni). Kada se procedura desezoniranja završi program u radnom fajlu sa podacima upisuje vrednosti sezonskog faktora, tj. ocenjene vrednosti sezonske komponente serije (**SAF_1**), vrednosti serije posle

⁵⁸ U novije vreme ciklična komponenta se najčešće tretira kao sastavni deo trenda i ne izdvaja se iz njega.

⁵⁹ Podroban prikaz modelskog pristupa desezoniranju može se naći u knjizi: Z. Kovačić (1998), Analiza vremenskih serija, Beograd: Ekonomski fakultet.

isključenja sezonske komponente (**SAS_1**), vrednosti za izravnatu trend-ciklus komponentu za vremensku serije iz koje je isključena sezonska komponenta (**STC_1**), kao i vrednosti za slučajnu, iregularnu komponentu serije (**ERR_1**). Ako je reč o multiplikativnom modelu dekompozicije vremenske serije tada iregularna, slučajna komponenta serije predstavlja količnik vrednosti vremenske serije iz koje je isključena sezonska komponenta (**SAS_1**) i izravnate trend-ciklus komponente vremenske serije posle serije iz koje je isključena sezonska komponenta (**STC_1**). U ispisu, se pored obaveštenja o imenima varijabli u kojima su sačuvani ovi podaci nalaze i sezonski indeksi po upotrebljenim sezonskim jedinicama (mesecima, kvartalima...).

Copyright @ Lazar Tenjović, 2005, sva prava zadržava