

6.2 Regresné prístupy k eliminovaniu sezónnej zložky

Predpokladajme existenciu sezónnej zložky. Najprv bude potrebné kvantifikovať sezónne výkyvy a tiež očistiť časový rad, t.j. eliminovať sezónnu zložku, ktorá zakrýva dynamiku ekonomických javov. Podoba modelov sezónnosti závisí na charaktere trendu a vzťahu trendovej a sezónnej zložky. Najčastejšie je tento problém riešený pomocou konštrukcie rozdielov a indexov alebo regresným prístupom k sezónnej zložke.

Pre elimináciu sezónnej zložky Sz_t môžeme použiť lineárne regresné metódy podobne, ako pri eliminácii trendu. Nevýhodou dekompozičnej analýzy časových radov so sezónnou zložkou je, že modely jednotlivých zložiek odhadujeme samostatne a následne určujeme aj prognózy jednotlivých zložiek.

Ak nezvolíme správne modely trendu a sezónnosti, chyby prognóz sa znásobia. V prípade aditívnej dekompozície (alebo po zlogaritmovaní pri multiplikatívnej dekompozícii) použijeme pomocné premenné nasledujúcim spôsobom (pre lineárny trend)

$$Sz_t = \alpha_2 x_{2t} + \dots + \alpha_L x_{Lt},$$

kde x_{2t}, \dots, x_{Lt} sú tzv. kvalitatívne premenné definované nasledujúcim spôsobom $x_{it} = 1$, ak čas t odpovedá i - tému obdobiu v roku, inak $x_{it} = 0$ (pre $i=2, \dots, L$).

Z hľadiska minimalizácie chýb prognóz sa preto odporúča zvoliť iba jeden model platný tak pre trendovú ako aj sezónnu zložku a odhad jeho parametrov získať simultánne. Vhodným nástrojom pre takéto modely časových radov sú modely viacnásobnej regresie s umelými premennými.

Pod umelou premennou sa rozumie nula - jednotková premenná, ktorá zastupuje jednotlivé sezóny tak, že v príslušnej sezóne jej priradíme jednotku a v ostatných sezónach obsahuje nuly.

Napríklad pre časové rady štvrťročných údajov so sezónnou zložkou, umelé premenné jednotlivých štvrťrokov sú (1,0,0,0) – pre 1. štvrťrok, (0,1,0,0) – pre 2. štvrťrok, (0,0,1,0) – pre 3. štvrťrok, (0,0,0,0) – pre 4. štvrťrok.

Teda

$$y_t = Tr_t + Sz_t + I_t, \text{ kde}$$

$$Tr_t = \beta_0 + \beta_1 t,$$

$$Sz_t = \alpha_2 x_{2t} + \dots + \alpha_L x_{Lt},$$

I_t je reziduálna zložka tvorená náhodnými fluktuáciami v priebehu časového radu, ktoré nemajú rozpoznateľný systematický charakter, pokrýva tiež chyby v meraniach a zaokrúhľovaniach.

Ak predpokladáme, že časový rad bude mať lineárny trend a štvrťročnú sezónnosť, aditívny model časového radu môžeme formulovať pomocou regresného modelu s umelými premennými v tvare

$$y_t = a_0 + a_1 t + b_1 Q_1 + b_2 Q_2 + b_3 Q_3 + I_t,$$

kde $Tr_t = a_0 + a_1 t$, $Sz_t = b_1 Q_1 + b_2 Q_2 + b_3 Q_3$. Sezónnu zložku zobrazujeme iba s $m-1=3$ umelými premennými. Premenná Q_4 má všetky hodnoty nulové, aby nebola vo vzájomnej závislosti s ostatnými umelými premennými.

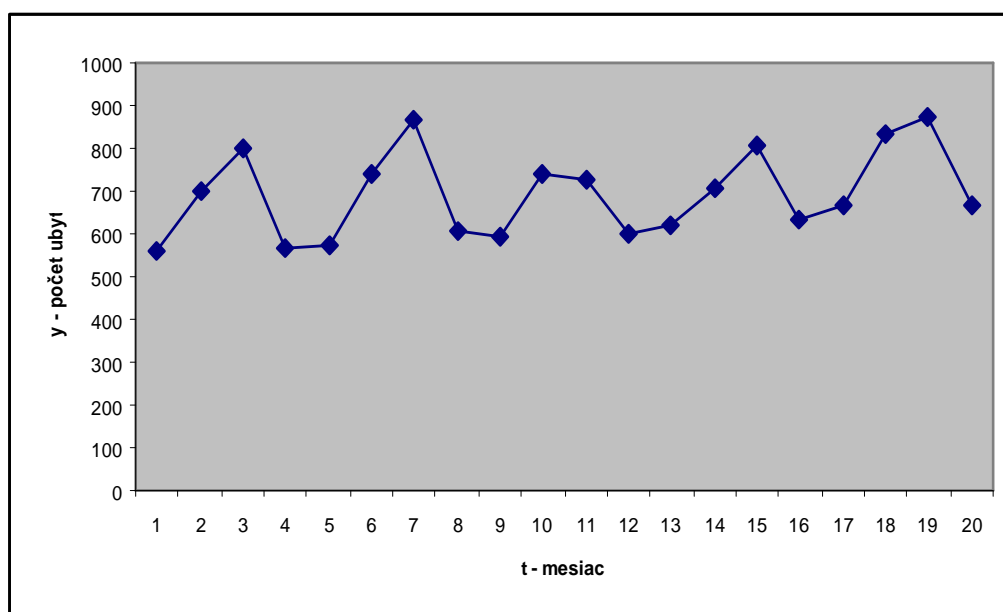
Príklad 6.1 V nasledujúcej tabuľke sú uvedené štvrťročné údaje o počte ubytovaných v rekreačnom zariadení XY v rokoch 1985 – 1989.

rok - počet ubytovaných	štvrťrok			
	1	2	3	4
1985	561	702	800	588
1986	575	738	868	605
1987	594	738	729	600
1988	622	708	806	632
1989	665	835	873	670

- Nakreslíme graf a posúdme či rad má trend a sezónnosť.
- Odhadnime trend.
- Určme sezónne indexy.
- Určme štvrťročné prognózy ubytovaných v jednotlivých štvrťrokoch v roku 1990.

Riešenie:

- Graf získaný pomocou Excelu: (vložiť graf – spojnicový, oblasť dát – počet ubytovaných)



Vidíme, že časový rad vykazuje sezónnosť.

- Trend určíme pomocou Excelu: *Nástroje - Analýza dát – Regresia*,
pole Y – označíme hodnoty y ,
pole X – označíme hodnoty t .

t	počet ubytovaných y_t	odhad trendu $Tr_t = 639,3689 + 5,2459t$	odhad $Sz_t = y_t / Tr_t$			
1	561	644,6148	0,870287			
2	702	649,8607	1,080232			
3	800	655,1066	1,221175			
4	568	660,3525	0,860147			
5	575	665,5984	0,863884			
6	738	670,8443	1,100106			
7	868	676,0902	1,283852			
8	605	681,3361	0,887961			
9	594	686,582	0,865155			
10	738	691,8279	1,066739			
11	729	697,0738	1,0458			
12	600	702,3197	0,854312			
13	622	707,5656	0,87907			
14	708	712,8115	0,99325			
15	806	718,0574	1,122473			
16	632	723,3033	0,873769			
17	665	728,5492	0,912773			
18	835	733,7951	1,13792			
19	873	739,041	1,181261			
20	670	744,2869	0,900191			
	Koeficienty	Chyba stř. hodnoty	t stat	Hodnota P	Dolní 95%	Horní 95%
Hranice	639,3684	46,1890	13,8424	0,0000	542,3289	736,4079
Soubor X 1	5,2459	3,8558	1,3605	0,1905	-2,8548	13,3466

Teda $Tr_t = 639,3684 + 5,2459t$

c) Sezónne indexy počtu ubytovaných

Rok	štvrt'rok			
	1.	2.	3.	4.
1985	0,870287	1,080232	1,221175	0,860147
1986	0,863884	1,100106	1,283852	0,887961
1987	0,865155	1,066739	1,0458	0,854312
1988	0,87907	0,99325	1,122473	0,873769
1989	0,912773	1,13792	1,181261	0,900191
Priemerné sezónne indexy $\hat{S}_{z,t}$	0,87823	1,07565	1,17091	0,87528

d) Prognózy počtu ubytovaných v roku 1990

rok	štvrťrok	T	sezónny index= \hat{S}_{z_t}	prognóza trendu $Tr_t = 639,3689 + 5,2459t$	prognóza počtu ubytovaných $Tr_t \cdot \hat{S}_{z_t}$
1990	1.	21	0,87823	749,5328	658,2622
	2.	22	1,07565	754,7787	811,8777
	3.	23	1,17091	760,0246	889,9204
	4.	24	0,87528	795,2705	669,8260

Teda v jednotlivých štvrtrokoch roku 1990 môžeme očakávať postupne 658; 812; 890; 670 ubytovaných.

Uvedieme iný prístup k riešeniu tohto príkladu.

Pretože sezónami sú jednotlivé štvrt'roky roka, vytvoríme postupne ich umelé premenné:

$$Q_1 = (1,0,0,0), Q_2 = (0,1,0,0), Q_3 = (0,0,1,0), Q_4 = (0,0,0,0)$$

[Chajdiak, J. - Rublíková, E. - Gudaba, M.: Štatistické metódy v praxi. STATIS Bratislava 1994, ISBN 80-85659-02-6]

Vieme, že trend vývoja je lineárny, takže celkový model budeme hľadať v tvare

$$y_t = B_0 + B_1 t + B_2 Q_1 + B_3 Q_2 + B_3 Q_3 + B_4 Q_4 + I_t,$$

pre $t = 1, 2, \dots, 20$. Parametre odhadneme pomocou MAPLEu.

```
>> with(stats):
```

```
fit[leastsquare[[t,q1,q2,q3,yt]]]([1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20],[1,0,0,0,1,0  
0,0,1,0,0,0,1,0,0,0,1,0,0,0],[0,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0],[0,0,1,0,0,0,1,0,0,0,1,0,0,1,0  
0,0,1,0],[561,702,800,568,575,738,868,605,594,738,729,600,622,708,806,632,665,835,873,670  
]);
```

$$y_t = \frac{11091}{20} + \frac{403t}{80} + \frac{281ql}{80} + \frac{5571q2}{40} + \frac{16419q3}{80}$$

Teda

$$\hat{y}_t = 554,55 + 5,0375t + 3,5125Q_1 + 139,275Q_2 + 205,2375Q_3.$$

Prognózy počtu ubytovaných na rok 1990 podľa tohto modelu (po zaokrúhlení) sú:

rok	T	štvrťrok	predpoved'
1990	21	1	663
	22	2	805
	23	3	876
	24	4	675