

9.4 Theilov koeficient

Ako miery prognostickej kvality modelu sa často používajú koeficienty nesúladu (nesúlad medzi simulovanou predpoveďou a v tom čase už známou skutočnosťou). Často používaný je **Theilov koeficient**

$$T_H^2 = \frac{\sum_{j=1}^D (y_{N+j} - \hat{P}_j)^2}{\sum_{j=1}^D (y_{N+j})^2},$$

kde N je dĺžka časového radu pre odhad modelu,

$D = n - N$, je skrátenie časového radu,

\hat{P}_j je extrapolácia (predpoveď) na j období dopredu a to modelom odhadnutým na základe prvých N pozorovaní.

V percentuálnom vyjadrení ($T_H \times 100\%$) môžeme túto veličinu interpretovať ako chybu predpovede vyjadrenej v percentách.

- Ak sa pohybuje v hraniciach 3 – 5% , potom je chyba predpovede považovaná za „malú“ a uvažovaný model môže byť vhodným nástrojom pre prognózovanie.
- Ak sa pohybuje v hraniciach 6 – 10% , potom použitie daného modelu nie je vylúčené.
- Ak je koeficient nesúladu väčší ako 10% , tak je potrebné zvážiť či nie je iný lepší model.

Príklad 9.1 V tabuľke sú údaje o počte áut v USA v rokoch 1960 až 1979. Určme koeficienty lineárneho trendu vývoja počtu áut použitím prvých 17 hodnôt, vypočítajme predpovede na roky 1977-79 a porovnajme ich so známymi údajmi. Mieru nesúladu simulovaných predpovedí so skutočnosťou vyhodnoťme Theilovým koeficientom.

rok x_i	t_i	počet áut (v mil.) y_i	$\hat{P}_i = Tr_i$	$(y_i - \hat{P}_i)^2$	y_i^2
1960	1	6,577			
1961	2	5,855			
1962	3	6,939			
1963	4	7,557			
1964	5	8,056			
1965	6	9,314			
1966	7	9,009			
1967	8	8,357			
1968	9	9,404			
1969	10	9,447			
1970	11	8,388			
1971	12	9,831			
1972	13	10,409			
1973	14	11,351			
1974	15	8,701			
1975	16	8,168			
1976	17	9,752			
1977	18	10,826	10,46141	0,132926597	117,202276
1978	19	10,946	10,66225	0,080512928	119,814916
1979	20	10,357	10,8631	0,256132149	107,267449
Suma				0,469571674	344,284641
a_0	6,846235				
a_1	0,200843				
TH^2	0,001364				

Riešenie:

Z prvých 17 hodnôt určíme trend pomocou metódy najmenších štvorcov. Dostávame

$$\hat{P}_i = Tr_i = a_0 + a_1 t_i = 6,846 + 0,200843 t_i.$$

Potrebné hodnoty pre $t = 18, 19, 20$ sú uvedené v tabuľke. Potom koeficient nesúlady je

$$T_H^2 = \frac{\sum_{j=1}^D (y_{N+j} - \hat{P}_j)^2}{\sum_{j=1}^D (y_{N+j})^2} = \frac{0,469571674}{344,284641} = 0,001364 ,$$

teda $T_H = 0,03693237$ čo odpovedá približne 3,7% .

Miera nesúlady nám potvrdzuje správnosť zvoleného lineárneho modelu.