

Trendy s konstantními parametry

Obecně lze trendové funkce lze rozdělit v podstatě do dvou skupin:

A. trendy lineární v parametrech (a nelineární pouze v proměnných)

V tomto případě lze trendovou složku psát ve tvaru **funkce lineární v parametrech**

$$(1) \quad Tr_t = y_t = \beta_0 + \beta_1 f_1(t) + \beta_2 f_2(t) + \beta_3 f_3(t) \dots + \beta_k f_k(t)$$

tzv. ve tvaru **lineárně-aditivního schématu** případně

$$(2) \quad Tr_t = g(y_t) = \beta_0 + \beta_1 f_1(t) + \beta_2 f_2(t) + \beta_3 f_3(t) \dots + \beta_k f_k(t) + \varepsilon_t$$

Vztah (2) se nazývá tvar **lineární v parametrech po transformaci vysvětlované proměnné** nebo také (podle W.Eichhorna) **zobecněná kvazilineární funkce**.¹

Výpočet parametrů trendových funkcí lineárních v parametrech (1) nečiní problém a provádí se **prostou metodou nejmenších čtverců MNČ/OLS**. Protože všechny vysvětlující proměnné jsou **nestochastické**, je situace s respektováním podmínek **standardního lineárního regresního modelu** usnadněna:

a) Přítomnost jedničkového vektoru zajišťuje nulovou střední hodnotu náhodných složek ε_t . (Konstantu β_0 lze interpretovat jako výchozí úroveň ukazatele v čase 0)

b) Je automaticky zajištěna **nestochastičnost** všech vysvětlujících proměnných (jde zpravidla o jednoduché transformace trendu). Není proto třeba ověřovat pro **konzistenci odhadové funkce** nutnou podmínku $Cov(x_{ij}, \varepsilon_t) = E(x_{ij} \cdot \varepsilon_t) = 0$, protože ta je zde splněna vždy.

c) V „matici plánu“, $[t, f_1(t), \dots, f_k(t)]$ se nemohou vyskytnout přesné lineární závislosti mezi „vysvětlujícími“, proměnnými (takže nemůže vzniknout **problém multikolinearity**)². Pokud nevolíme dále funkce f_j příliš „blízké“ (např. $t^{17/18}$ a $t^{18/19}$), nehrozí ani **multikolinearita přibližná**.

V obou předchozích případech přijímáme aditivní připojení náhodné složky ε_t , takže stochastický/regresní tvar trendové (1) funkce je tedy

$$(1St) \quad y_t = \beta_0 + \beta_1 f_1(t) + \beta_2 f_2(t) + \beta_3 f_3(t) \dots + \beta_k f_k(t) + \varepsilon_t \text{ a obdobně}$$

Stochastický/regresní tvar trendové (2) závislosti je

$$(2St) \quad g(y_t) = \beta_0 + \beta_1 f_1(t) + \beta_2 f_2(t) + \beta_3 f_3(t) \dots + \beta_k f_k(t) + \varepsilon_t .$$

¹ **Kvazilineární funkce** může být zapsána schématem

$$F(x_1, \dots, x_s, \dots, x_n) = g^{-1}(a + b_1 g(x_1) + \dots + b_n g(x_n)) , \quad \text{kde}$$

g je spojitá monotónní funkce a dále a, b_1, \dots, b_n jsou vhodné konstanty (a nenulové).

Zobecněnou kvazilineární funkci F lze zapsat ve tvaru:

$$F(x_1, \dots, x_s, \dots, x_n) = h(f_1(x_1), \dots, f_s(x_s), \dots, f_n(x_n)) , \quad \text{kde}$$

f_1, \dots, f_n jsou vesměs spojitě a ryze monotónní funkce. Oproti **kvazilineární funkci** se nevyžaduje symetrie vnitřních funkcí $g_i, i = 1, \dots, n$ vůči jednotlivým argumentům.

² S přirozenou podmínkou, že funkce $f_1(t), \dots, f_k(t)$ jsou vzájemně různé.

Standardní OLS-kritérium pro výpočet trendových parametrů má tvar

pro případ (1St)
$$\text{Min} \sum_{t=1}^n [y_t - \hat{y}_t]^2 = \text{Min} \sum_{t=1}^n \left[y_t - \sum_{j=1}^k \beta_j f_j(t) \right]^2$$
 přes $\beta_1, \beta_2, \dots, \beta_k$, ale

pro případ (2St)
$$\text{Min} \sum_{t=1}^n [g(y_t) - g(\hat{y}_t)]^2 = \text{Min} \sum_{t=1}^n \left[g(y_t) - g\left(\sum_{j=1}^k \beta_j f_j(t) \right) \right]^2$$

V jednom i ve druhém případě lze užít *prostou metodou nejmenších čtverců MNČ/OLS* bez korekcí, avšak za uvědomění stojí, že v případě (4) OLS-kritérium neminimalizuje rozdíl prostých, nýbrž transformovaných (funkcí g) pozorovaných y_t a vyrovnaných \hat{y}_t hodnot časové řady.³

B. trendy nelineární v parametrech (*případně nelineární i v proměnných*) .

V tomto obecném případě lze trendovou funkci psát ve tvaru

(3)
$$Tr_t = y_t = h(\beta_0, \beta_1, \beta_2, \beta_3, \dots, \beta_k, f_1(t), f_2(t), \dots, f_r(t))$$

případně ještě obecněji

(4)
$$Tr_t = g(y_t) = h(\beta_0, \beta_1, \beta_2, \beta_3, \dots, \beta_k, t) .$$

Odpovídající stochastická specifikace tvar trendové funkce (3) je následně

(3St)
$$y_t = h(\beta_0, \beta_1, \beta_2, \beta_3, \dots, \beta_k, f_1(t), f_2(t), \dots, f_r(t)) + \varepsilon_t$$
 a obdobně

Stochastický/regresní tvar trendové závislosti (4) je pak

(4St)
$$Tr_t = g(y_t) = h(\beta_0, \beta_1, \beta_2, \beta_3, \dots, \beta_k, t) + \varepsilon_t .$$

V tomto případě již nelze k výpočtu parametrů (tzn. k jejich konzistentnímu odhadu) použít standardní *OLS-odhadovou funkci*. K minimalizaci výrazu (4) je zapotřebí (s výjimkou ojedinělých příznivých podob nelinearity) uplatnit některou z metod nelineární optimalizace, jako je *NLLS (nelineární metoda nejmenších čtverců)* nebo *NLLAD (nelineární metoda nejmenších absolutních odchylek)* s kritérii

pro specifikaci (3St) a *NLLS*

$$\text{Min} \sum_{t=1}^n e_t^2 = \text{Min} \sum_{t=1}^n [y_t - \hat{y}_t]^2 = \text{Min} \sum_{t=1}^n [y_t - \hat{h}(\beta_0, \beta_1, \beta_2, \dots, \beta_k, t)]^2$$
 přes $\beta_1, \beta_2, \dots, \beta_k$

resp. pro specifikaci (3St) a *NLLAD*

$$\text{Min} \sum_{t=1}^n |e_t| = \text{Min} \sum_{t=1}^n |y_t - \hat{y}_t| = \text{Min} \sum_{t=1}^n |y_t - \hat{h}(\beta_0, \beta_1, \beta_2, \dots, \beta_k, t)|$$
 přes $\beta_1, \beta_2, \dots, \beta_k$

Poznámka: Nejobecnější myslitelný tvar trendové nelineární specifikace by byl

(4St)
$$G(\beta_0, \beta_1, \beta_2, \beta_3, \dots, \beta_k, f_1(t), f_2(t), \dots, f_r(t), \varepsilon_t) = 0 .$$

Ten je však natolik neurčitý a výpočetně problematický, že se neuzívá.

³ Odhady parametrů získané pomocí jednoho a druhého kritéria tedy nebudou obecně shodné.

Je patrné, že **konzistentní odhad trendových parametrů dostaneme** v prvním případě nasazením **proste metody nejmenších čtverců OLS**, zatímco ve druhém je nutné uplatnit **nelineární prostou metodu nejmenších čtverců NLLS**.

Alternativním postupem pro výpočet trendových parametrů může dále být např. **metoda nejmenších absolutních odchylek LAD**, jejíž minimalizační kritérium je

pro případ (3St)
$$\text{Min} \sum_{t=1}^n |y_t - \hat{y}_t| = \text{Min} \sum_{t=1}^n \left| y_t - \sum_{j=1}^k \beta_j f_j(t) \right|$$

pro případ (4St)
$$\text{Min} \sum_{t=1}^n |g(y_t) - g(\hat{y}_t)| = \text{Min} \sum_{t=1}^n \left| g(y_t) - g \left(\sum_{j=1}^k \beta_j f_j(t) \right) \right|$$

Pravidla pro výpočty diferencí

1.diference: $d_1 = y_2 - y_1$, $d_2 = y_3 - y_2$, $d_3 = y_4 - y_3$, $d_4 = y_5 - y_4$, atd

obecně: $\Delta y_t = d_t = y_{t+1} - y_t$

2.diference: $e_1 = d_2 - d_1 = y_3 - 2y_2 + y_1$, $e_2 = d_3 - d_2 = y_4 - 2y_3 + y_2$

obecně: $\Delta^2 y_t = e_t = y_{t+2} - 2y_{t+1} + y_t$

3.diference: $f_1 = e_2 - e_1 = y_4 - 3y_3 + 3y_2 - y_1$, $f_2 = e_3 - e_2 = y_5 - 3y_4 + 3y_3 - y_2$

obecně: $\Delta^3 y_t = f_t = y_{t+3} - 3y_{t+2} + 3y_{t+1} - y_t$

4.diference: $g_1 = f_2 - f_1 = y_5 - 4y_4 + 6y_3 - 4y_2 + y_1$

obecně: $\Delta^4 y_t = g_t = y_{t+4} - 4y_{t+3} + 6y_{t+2} - 4y_{t+1} + y_t$

Jak patrně, **sestava koeficientů odpovídá analogickému schématu pro binomické součinitele u mnohočlenu typu** $P(a,b) = (a-b)^k$

Nejpoužívanější matematické funkce při trendovém vyrovnání/extrapolaci

1. Lineární trend (11) $y_t = \beta_0 + \beta_1 t$

indikace přítomnosti: první diference $y_{t+1} - y_t$ jsou přibližně konstantní.

(Trend je monotónní, zda rostoucí či klesající určuje znaménko koeficientu β_1).

2. Kvadratický trend (12) $y_t = \beta_0 + \beta_1 t + \beta_2 t^2$

indikace přítomnosti: druhé diference $y_{t+2} - 2y_{t+1} + y_t$ jsou zhruba konstantní

(Trend není monotónní, chování v „nekonečné budoucnosti“ $+\infty$ určuje znaménko koeficientu β_2).

3. Kubický trend (13) $y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3$

indikace přítomnosti: třetí diference $y_{t+3} - 3y_{t+2} + 3y_{t+1} - y_t$ jsou zhruba konstantní

(Trend není monotónní, má nejvýš dvě „úvratě“: tzn. jedno lokální maximum, jedno minimum) chování v „nekonečné budoucnosti“ $+\infty$ určuje znaménko koeficientu β_3).

4. Polynomický trend s-tého stupně $y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \dots + \beta_s t^s$

indikace přítomnosti: diference s-tého stupně jsou přibližně konstantní

5 Logaritmický trend (15A) $y_t = \beta_0 + \beta_1 \ln(t)$ (15B) $y_t = \beta_0 + \beta_1 \log(t)$

(Trend je monotónní, růst (typické chování) či klesání určuje znaménko koeficientu β_1).

indikace přítomnosti: $y_{t+1} = \beta_0 + \beta_1 \ln(t+1)$

$$y_{t+1} - y_t = \beta_1 \ln(t+1) - \beta_1 \ln(t) = \beta_1 \ln \frac{t+1}{t} = \beta_1 \ln \left(1 + \frac{1}{t} \right)$$

$$e^{y_{t+1}} - e^{y_t} = \beta_1 \ln(t+1) - \beta_1 \ln(t) = \beta_1 \ln \frac{t+1}{t} \text{ závisí na } t.$$

Diference „expomocnin,, konvergují s rostoucím t k nule

6. Mocninový trend (16) $y_t = \beta_0 t^{\beta_1}$, $\beta_0 > 0$, β_1 zpravidla $\beta_1 \in (0, 2)$

(Monotónní trend, intenzitu růstu určuje koeficient β_1 : při $\beta_1 > 1$ konvexní, při $\beta_1 < 1$ konkávní průběh)

indikace přítomnosti: a) vyjádřeme $y_t = \beta_0 t^{\beta_1}$, $y_{t+1} = \beta_0 (t+1)^{\beta_1}$

Podíly původních hodnot: $\frac{y_{t+1}}{y_t} = \frac{\beta_0 (t+1)^{\beta_1}}{\beta_0 t^{\beta_1}} = \left(\frac{t+1}{t} \right)^{\beta_1} = \left(1 + \frac{1}{t} \right)^{\beta_1}$ konvergují k 1

b) vyjádřeme $\ln y_t = \ln \beta_0 + \beta_1 \ln(t)$ $\ln y_{t+1} = \ln \beta_0 + \beta_1 \ln(t+1)$

$$\ln y_{t+1} - \ln y_t = \beta_1 \ln(t+1) - \beta_1 \ln(t) = \beta_1 \ln \frac{t+1}{t} = \beta_1 \ln \left(1 + \frac{1}{t} \right)$$

Diference logaritmovaných hodnot:

konvergují s rostoucím t k 0, (ale závisí na t)

7. Exponenciální trend (17) $y_t = \beta_0 \cdot \beta_1^t$ [$\beta_0 > 0$, zpravidla $\beta > 1$]

indikace přítomnosti: podíly sousedních hodnot y_{t+1} / y_t resp. první diference logaritmů $\ln y_{t+1} - \ln y_t$ jsou přibližně konstantní.

(Monotónní trend, intenzitu růstu určuje koeficient β_1 : při $\beta_1 > 1$ konvexní, při $\beta_1 < 1$ konkávní průběh).

8. Modifikovaný exponenciální trend (18) $y_t = \gamma + \alpha\beta^t$

indikace přítomnosti: podíly sousedních prvních diferencí $(y_{t+2} - y_{t+1}) / (y_{t+1} - y_t)$ jsou přibližně konstantní

9. Logistický trend (19) $y_t = \frac{\gamma}{1 + \alpha\beta^t}$ s parametry $\beta > 0, \gamma > 0$

indikace přítomnosti: a) histogram prvních diferencí $y_{t+1} - y_t$ je tvarem podobná křivce (hustotě) normovaného normálního rozdělení $N(0,1)$.

b) podíly sousedních prvních diferencí reciprokových hodnot $(1/y_{t+2} - 1/y_{t+1}) / (1/y_{t+1} - 1/y_t)$ jsou přibližně konstantní.

Inflexní bod logistické křivky je v bodě $t^* = -\frac{\ln(\alpha)}{\ln(\beta)}$.

10. Gompertzův trend $\ln(y_t) = \gamma + \alpha\beta^t$ nebo ekvivalentně (20) $y_t = \exp(\gamma + \alpha\beta^t)$ [$\beta > 0$]

indikace přítomnosti: podíly prvních logaritmovaných diferencí $(\ln y_{t+2} - \ln y_{t+1}) / (\ln y_{t+1} - \ln y_t)$ jsou přibližně konstantní

Inflexní bod má Gompertzova křivka v bodě $t^* = -\frac{\ln(-\alpha)}{\ln(\beta)}$

11. Odmocninný trend $y_t = \beta_0 + \beta_1\sqrt{t}$

12. Hyperbolický trend $y_t = \beta_0 + \frac{\beta_1}{t}$

13. Lineárně-hyperbolický trend $y_t = \beta_0 + \frac{\beta_1}{t}$

1A) výpočet parametrů lineárního trendu $Tr^{LIN}_t = \beta_0 + \beta_1 t$

Pro jejich určení lze nejnázorněji uplatnit výraz pro **OLS-minimalizační kritérium** z výchozí regresní specifikace (3)

$$\text{Min} \sum_{t=1}^n [y_t - \hat{y}_t]^2 = \text{Min} \sum_{t=1}^n \left[y_t - \sum_{j=1}^2 \beta_j f_j(t) \right]^2, \text{ zde konkretizované na}$$

$$(21) \quad \text{Min} \sum_{t=1}^n [y_t - \hat{y}_t]^2 = \text{Min} \sum_{t=1}^n [y_t - \beta_0 - \beta_1 t]^2.$$

a) Uplatněním standardního postupu – derivováním (21) podle obou neznámých parametrů dostaneme soustavu dvou nutných podmínek extrému

$$(22A,B) \quad 2 \sum_{t=1}^n [y_t - \beta_0 - \beta_1 t](-1) = 0, \text{ resp. } 2 \sum_{t=1}^n [y_t - \beta_0 - \beta_1 t](-t) = 0$$

a po očividných úpravách

$$\sum_{t=1}^n [y_t - \beta_0 - \beta_1 t] = 0, \text{ resp. } \sum_{t=1}^n [y_t - \beta_0 - \beta_1 t]t = 0 \text{ neboli}$$

$$(23A,B) \quad \sum_{t=1}^n y_t - \beta_0 \cdot n - \beta_1 \sum_{t=1}^n t = 0, \text{ resp. } \sum_{t=1}^n t y_t - \beta_0 \cdot \sum_{t=1}^n t - \beta_1 \sum_{t=1}^n t^2 = 0$$

Přeskupením výrazů na obou stranách dospějeme k soustavě **normálních rovnic** pro odhadnuté parametry $\hat{\beta}_0, \hat{\beta}_1$ - zde značené b_0, b_1 - ve tvaru

$$(24A) \quad b_0 \cdot n + b_1 \sum_{t=1}^n t = \sum_{t=1}^n y_t$$

$$(24B) \quad b_0 \cdot \sum_{t=1}^n t + b_1 \sum_{t=1}^n t^2 = \sum_{t=1}^n t y_t, \text{ jejímž řešením je dvojice}$$

$$(25A,B) \quad b_0 = \frac{\sum_{t=1}^n t^2 \cdot \sum_{t=1}^n y_t - \sum_{t=1}^n t \cdot \sum_{t=1}^n t y_t}{n \cdot \sum_{t=1}^n t^2 - \left(\sum_{t=1}^n t \right)^2} \quad b_1 = \frac{n \cdot \sum_{t=1}^n t y_t - \sum_{t=1}^n t \cdot \sum_{t=1}^n y_t}{n \cdot \sum_{t=1}^n t^2 - \left(\sum_{t=1}^n t \right)^2}$$

Po úpravách součtových výrazů $\sum_{t=1}^n t = \frac{n(n+1)}{2}$, $\sum_{t=1}^n t^2 = \frac{n(n+1)(2n+1)}{6}$ a vydělením čitatele i jmenovatel n dostaneme pro b_1

$$(26) \quad b_1 = \frac{\sum_{t=1}^n t y_t - \frac{(n+1)}{2} \cdot \sum_{t=1}^n y_t}{\frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)^2}{4}} = \frac{\sum_{t=1}^n t y_t - \frac{(n+1)}{2} \cdot \sum_{t=1}^n y_t}{\frac{n(n+1)[4n+2-3(n+1)]}{12}} = \frac{\sum_{t=1}^n t y_t - \frac{(n+1)}{2} \cdot \sum_{t=1}^n y_t}{\frac{n(n^2-1)}{12}}$$

a následně pro b_0 $b_0 = \bar{y} - b_1 \bar{t} = \bar{y} - \frac{n+1}{2} b_1$

b) Ke stejným výrazům dospějeme aplikací standardního vzorce pro **OLS-odhady parametrů**, tedy $\hat{\beta} = (X'X)^{-1} X'y$, pokud v něm konkretizujeme matici X jako

$$X = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ \dots & \dots \\ 1 & n \end{pmatrix}. \text{ Potom máme } \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} n & \sum_{t=1}^n t \\ \sum_{t=1}^n t & \sum_{t=1}^n t^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{t=1}^n y_t \\ \sum_{t=1}^n ty_t \end{pmatrix}$$

Zřejmě máme $(X'X)^{-1} = \frac{1}{n \sum_{t=1}^n t^2 - \left(\sum_{t=1}^n t\right)^2} \begin{pmatrix} \sum_{t=1}^n t^2 & -\sum_{t=1}^n t \\ -\sum_{t=1}^n t & n \end{pmatrix}$, $X'y = \begin{pmatrix} \sum_{t=1}^n y_t \\ \sum_{t=1}^n ty_t \end{pmatrix}$.

Výpočtem (řešením pro neznámé b_0, b_1) dostaneme dvojici odhadů (25A,B). Že jde skutečně o minimum, se snadno přesvědčíme stejně snadno jako v případě minimalizace SSE (součtu čtverců reziduí) v lineárním regresním modelu.

Předpověď y_T * **budoucí hodnoty** y_T má u lineárního trendu tvar

(27) $y^*_T = b_0 + b_1 T$.

Příslušný $100 \cdot (1 - p)$ procentní předpovědní interval (při neznalosti σ_e) je

(28) $(\hat{y}_T - t_{1-p/2}(n-2) \cdot s_e f_T; \hat{y}_T + t_{1-p/2}(n-2) \cdot s_e f_T)$, kde

\hat{y}_t je vyrovnaná hodnota závisle proměnné

$n - 2$ je počet stupňů volnosti (rozdíl mezi počtem pozorování a počtem trendových parametrů)

$t_{1-p/2}$ je $1-p/2$ (.100) % kvantil Studentova t-rozdělení o $n-2$ stupních volnosti

f_T je hodnota: $f_T = \sqrt{1 + (1, T)(X'X)^{-1}(1, T)}$, přičemž

$$X = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ \dots & \dots \\ 1 & n \end{pmatrix} \text{ je matice nestochastických „regresorů“}.$$

S ohledem na specifický výraz pro momentovou matici v daném případě

$$(X'X)^{-1} = \frac{1}{\frac{(n+1)(2n+1)}{12} - \frac{n^2(n+1)^2}{4}} \begin{pmatrix} \frac{n(n+1)(2n+1)}{12} & -\frac{n(n+1)}{2} \\ -\frac{n(n+1)}{2} & n \end{pmatrix}$$

dostaneme (29)

$$f_T = \sqrt{1 + \frac{1}{n} + \frac{\left(T - \frac{n+1}{2}\right)^2}{n(n^2 - 1) \cdot 12}}$$

s_e je směrodatná odchylka reziduí (rozdílu mezi pozorováními a vyrovnanými

hodnotami časové řady) (30)

$$s_e = \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n-2}} = \sqrt{\frac{\sum_{t=1}^n y_t^2 - \sum_{t=1}^n \hat{y}_t^2}{n-2}}$$

2A) výpočet parametrů kvadratického trendu $Tr^Q_t = \beta_0 + \beta_1 t + \beta_2 t^2$

Aplikace standardního vzorce pro OLS $\hat{\beta} = (X'X)^{-1} X'y$ vede k výrazům

$$(31) \quad \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} n & \sum_{t=1}^n t & \sum_{t=1}^n t^2 \\ \sum_{t=1}^n t & \sum_{t=1}^n t^2 & \sum_{t=1}^n t^3 \\ \sum_{t=1}^n t^2 & \sum_{t=1}^n t^3 & \sum_{t=1}^n t^4 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{t=1}^n y_t \\ \sum_{t=1}^n t y_t \\ \sum_{t=1}^n t^2 y_t \end{pmatrix}$$

Odtud vyplývá soustava normálních rovnic ve tvaru

$$(32A,B,C) \quad \begin{aligned} b_0 \cdot n + b_1 \sum t + b_2 \sum t^2 &= \sum y_t \\ b_0 \cdot \sum t + b_1 \sum t^2 + b_2 \sum t^3 &= \sum t y_t \\ b_0 \cdot \sum t^2 + b_1 \sum t^3 + b_2 \sum t^4 &= \sum t^2 y_t \end{aligned}$$

Někdy je výhodné pracovat s vyjádřením trendu ve tvaru

$$(33) \quad Tr_t = \beta_0 + \beta_1(t - \bar{t}) + \beta_2(t - \bar{t})^2 \quad t = 1, 2, \dots, n, \text{ kde } \bar{t} = \frac{1}{n} \sum_{t=1}^n t = \frac{n+1}{2}$$

Protože zde platí $\sum_{t=1}^n (t - \bar{t}) = \sum_{t=1}^n (t - \bar{t})^3 = 0 \quad t = 1, 2, \dots, n$.

Předpověď y_T^* * **budoucí hodnoty** y_T má tvar

$$(34) \quad y^*_T = b_0 + b_1 T + b_2 T^2$$

Příslušný $100 \cdot (1 - p)$ **procentní předpovědní interval** je

$$(35) \quad (y_T^* - t_{1-p/2}(n-3) \cdot s_e f_T ; y_T^* + t_{1-p/2}(n-3) \cdot s_e f_T)$$
 , kde

\hat{y}_t je vyrovnaná hodnota závisle proměnné

$n - 3$ je počet stupňů volnosti (rozdíl mezi počtem pozorování a počtem trendových parametrů)

$t_{1-p/2}(n-3)$ je $100 \cdot \left(1 - \frac{p}{2}\right)\%$ - ní kvantil **Studentova t-rozdělení** $n-3$ stupních volnosti

f_T je hodnota: (36) $f_T = \sqrt{1 + (1, T, T^2)(X'X)^{-1}(1, T, T^2)}$, přičemž

$$X = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ 1 & n & n^2 \end{pmatrix} \text{ je matice nestochastických „regresorů“ .}$$

s_e je směrodatná odchylka reziduí (rozdílu mezi pozorovanými a vyrovnanými

hodnotami časové řady) (37) $s_e = \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n-3}}$

5B) výpočet parametrů logaritmického trendu $y_t^{Lo} = \beta_0 + \beta_1 \cdot \ln(t)$

Aplikace standardního vzorce pro OLS $\hat{\beta} = (X'X)^{-1} X'y$ má zde podobu (42)

$$\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} n & \sum_{t=1}^n \ln t \\ \sum_{t=1}^n \ln t & \sum_{t=1}^n \ln^2 t \end{pmatrix}^{-1} \cdot \begin{pmatrix} \sum_{t=1}^n y_t \\ \sum_{t=1}^n y_t \ln t \end{pmatrix} = \frac{1}{n \sum_{t=1}^n \ln^2 t - \left(\sum_{t=1}^n \ln t \right)^2} \begin{pmatrix} \sum_{t=1}^n \ln^2 t & - \sum_{t=1}^n \ln t \\ - \sum_{t=1}^n \ln t & n \end{pmatrix} \begin{pmatrix} \sum_{t=1}^n y_t \\ \sum_{t=1}^n y_t \ln t \end{pmatrix},$$

takže pro parametrické odhady (budou konzistentní a nestranné) máme

$$b_0 = \bar{\beta}_0 = \frac{\left(\sum_{t=1}^n \ln^2 t \right) \left(\sum_{t=1}^n y_t \right) - \left(\sum_{t=1}^n \ln t \right) \left(\sum_{t=1}^n y_t \ln t \right)}{n \sum_{t=1}^n \ln^2 t - \left(\sum_{t=1}^n \ln t \right)^2}, \quad b_1 = \bar{\beta}_1 = \frac{n \left(\sum_{t=1}^n y_t \ln t \right) - \left(\sum_{t=1}^n \ln t \right) \left(\sum_{t=1}^n y_t \right)}{n \sum_{t=1}^n \ln^2 t - \left(\sum_{t=1}^n \ln t \right)^2},$$

6B) výpočet parametrů mocninného trendu

Zlogaritmováním definičního výrazu $y_t = \beta_0 t^{\beta_1}$ dostaneme $\ln y_t = \ln \beta_0 + \beta_1 \cdot \ln(t)$

Konkretizace standardního vzorce pro OLS $\hat{\beta} = (X'X)^{-1} X'y$ má zde podobu (43)

$$\begin{pmatrix} \ln \bar{\beta}_0 \\ \bar{\beta}_1 \end{pmatrix} = \begin{pmatrix} n & \sum_{t=1}^n \ln t \\ \sum_{t=1}^n \ln t & \sum_{t=1}^n \ln^2 t \end{pmatrix}^{-1} \cdot \begin{pmatrix} \sum_{t=1}^n \ln y_t \\ \sum_{t=1}^n \ln y_t \ln t \end{pmatrix} = \frac{1}{n \sum_{t=1}^n \ln^2 t - \left(\sum_{t=1}^n \ln t \right)^2} \begin{pmatrix} \sum_{t=1}^n \ln^2 t & - \sum_{t=1}^n \ln t \\ - \sum_{t=1}^n \ln t & n \end{pmatrix} \begin{pmatrix} \sum_{t=1}^n \ln y_t \\ \sum_{t=1}^n \ln y_t \ln t \end{pmatrix}$$

+takže konvenční OLS-regresí dostaneme konzistentní odhady pro $\ln \beta_0$ a β_1 . Vyexponováním prvního z obou odhadů pak získáme $\bar{\beta}_0$ jako $\exp(\ln \bar{\beta}_0)$, druhý obdržíme přímo. Poznamenejme ale současně, že v tomto případě nejde exaktně o OLS-odhady, neboť předmětem optimalizace

není výraz $\text{Min} \sum_{t=1}^T [y_t - \beta_0 t^{\beta_1}]^2$, nýbrž výraz $\text{Min} \sum_{t=1}^T [\ln y_t - \ln \beta_0 - \beta_1 \cdot \ln(t)]^2$.

Pro získání přesného odhadu bychom museli na první z obou minimandů uplatnit nelineární metodu nejmenších čtverců NLLS (zde ovšem nelze výsledný výraz zapsat explicitní formou). Viz obecnější poznámka níže.

11B) výpočet parametrů odmocninného trendu $y_t = \beta_0 + \beta_1 \cdot \sqrt{t}$

Aplikuje-li standardní **OLS – odhadové schéma** pro tento model, který je lineární v parametrech, dostaneme

$$(91) \quad \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} n & \sum_{t=1}^n \sqrt{t} \\ \sum_{t=1}^n \sqrt{t} & \sum_{t=1}^n t \end{pmatrix}^{-1} \begin{pmatrix} \sum_{t=1}^n y_t \\ \sum_{t=1}^n \sqrt{t} \cdot y_t \end{pmatrix}$$

$$\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} n & \sum_{t=1}^n \sqrt{t} \\ \sum_{t=1}^n \sqrt{t} & \sum_{t=1}^n t \end{pmatrix}^{-1} \begin{pmatrix} \sum_{t=1}^n y_t \\ \sum_{t=1}^n y_t \sqrt{t} \end{pmatrix} = \frac{1}{\frac{n^2(n+1)}{2} - \left(\sum_{t=1}^n \sqrt{t}\right)^2} \begin{pmatrix} \frac{n(n+1)}{2} & -\sum_{t=1}^n \sqrt{t} \\ -\sum_{t=1}^n \sqrt{t} & n \end{pmatrix} \begin{pmatrix} \sum_{t=1}^n y_t \\ \sum_{t=1}^n y_t \sqrt{t} \end{pmatrix},$$

takže pro parametrické odhady (budou konzistentní a nestranné) máme

$$b_0 = \hat{\beta}_0 = \frac{\frac{n(n+1)}{2} \left(\sum_{t=1}^n y_t\right) - \left(\sum_{t=1}^n \sqrt{t}\right) \left(\sum_{t=1}^n y_t \sqrt{t}\right)}{\frac{n^2(n+1)}{2} - \left(\sum_{t=1}^n \sqrt{t}\right)^2}, \quad b_1 = \hat{\beta}_1 = \frac{n \left(\sum_{t=1}^n y_t \sqrt{t}\right) - \left(\sum_{t=1}^n \sqrt{t}\right) \left(\sum_{t=1}^n y_t\right)}{\frac{n^2(n+1)}{2} - \left(\sum_{t=1}^n \sqrt{t}\right)^2},$$

13B) výpočet parametrů lineárně hyperbolického trendu

$$Tr^{Hy}_t = \beta_0 + \beta_1 t + \beta_2 \frac{1}{t}$$

Aplikace standardního vzorce pro OLS $\hat{\beta} = (X'X)^{-1} X'y$ vede k výrazům

$$(41) \quad \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} n & \sum_{t=1}^n t & \sum_{t=1}^n \frac{1}{t} \\ \sum_{t=1}^n t & \sum_{t=1}^n t^2 & \sum_{t=1}^n t \\ \sum_{t=1}^n \frac{1}{t} & \sum_{t=1}^n t & \sum_{t=1}^n \frac{1}{t^2} \end{pmatrix}^{-1} \begin{pmatrix} \sum_{t=1}^n y_t \\ \sum_{t=1}^n t \cdot y_t \\ \sum_{t=1}^n \frac{y_t}{t} \end{pmatrix}$$

Poznámka V případech, kdy je levostranná regresní proměnná transformací původní veličiny časové řady y_t , lze aplikovat metodu nejmenších čtverců OLS jen s výhradou.

Minimalizace reziduí prováděná pomocí OLS nebude totiž založena na kritériu $Min e_t^2$, nýbrž na kritériu $Min [f(e_t)]^2$, kde $f(\cdot)$ je použitá transformující funkce, nejčastěji logaritmická funkce nebo exponenciála. Pro takovéto případy by měla být užita nelineární (prostá) metoda nejmenších čtverců **NLLS (Non-Linear Least Squares Method)**, jejímž minimalizačním kritériem je právě

$$Min \sum_{t=1}^n [f(y_t) - f(\hat{y}_t)]^2$$

Výpočet je ale nutné provést některou z **numerických iteračních metod**, protože výsledný vzorec pro odhadnuté parametry nelze vyjádřit explicitně.

5) výpočet parametrů exponenciálního trendu

$$(51) \quad Tr^{ex}_t = \beta_0 \beta_1^t \quad [\beta_0 > 0] \text{ a obvykle } [\beta_1 > 1]$$

Trend je charakteristický tím, že **koeficient růstu** tj. **podíl dvou sousedních hodnot** T_{t+1}/T_t a současně podíl dvou sousedních diferencí $(T_{t+2} - T_{t+1})/(T_{t+1} - T_t)$

má konstantní hodnotu β_1 . Parametr β_0 se uvažuje (vždy) kladný.

Růstová tendence nastává v případě $\beta_1 > 1$, tendence poklesu v případě $\beta_1 < 1$,

K výpočtu parametrů exponenciálního trendu je vhodné uplatnit **váženou metodu nejmenších čtverců WLS** s vhodně transformovanými vahami. (Nelze totiž předpokládat multiplikativní tvar a logaritmicko-normální rozdělení náhodné složky v původním modelu před transformací).

V konkrétním případě exponenciálního trendu se aplikuje **WLS** pro minimalizaci výrazu

$$(52) \quad \sum_{t=1}^n v_t (y_t - \beta_0 \beta_1^t)^2, \text{ v němž } v_t \text{ jsou předem zvolené váhy.}$$

Minimalizace se však v této metodě vztahuje k výrazu

$$(53) \quad \sum_{t=1}^n w_t (\ln y_t - \ln \beta_0 - t \ln \beta_1)^2, \text{ u kterého}$$

váhy w_t závisí na vahách v_t tak, abychom minimalizací obou výrazů (52),(53) dostali aspoň přibližně shodné odhady parametrů b_0, b_1 . Ukazuje se, že pro přijatou logaritmickou transformaci je vhodné položit

$$(54) \quad w_t = y_t^2 \cdot v_t \quad t = 1, 2, 3, \dots, n$$

K nejčastější volbě vah $v_t = 1$ přistupujeme tehdy, jestliže nemáme důvod preferovat v minimalizačním kritériu některá jednotlivá pozorování, takže pak pro transformované váhy bude platit $w_t = y_t^2$. Minimalizací výrazu (53) s těmito vahami obdržíme **soustavu normálních rovnic**

$$(55A) \quad \ln b_0 \cdot (\sum y_t^2) + \ln b_1 (\sum t \cdot y_t^2) = (\sum y_t^2 \cdot \ln y_t)$$

$$(55B) \quad \ln b_0 \cdot (\sum t \cdot y_t^2) + \ln b_1 (\sum t^2 \cdot y_t^2) = (\sum t y_t^2 \cdot \ln y_t),$$

neboť po dosazení (54) do (53) máme

$$(53^*) \quad \sum_{t=1}^n y_t^2 (\ln y_t - \ln \beta_0 - t \ln \beta_1)^2, \text{ a tedy}$$

$$(56A) \quad \frac{\partial \sum_{t=1}^n y_t^2 (\ln y_t - \ln \beta_0 - t \ln \beta_1)^2}{\partial \beta_0} = -2 \cdot \sum_{t=1}^n y_t^2 (\ln y_t - \ln \beta_0 - t \ln \beta_1) \cdot \frac{1}{\beta_0}$$

$$(56B) \quad \frac{\partial \sum_{t=1}^n y_t^2 (\ln y_t - \ln \beta_0 - t \cdot \ln \beta_1)^2}{\partial \beta_1} = -2 \cdot \sum_{t=1}^n y_t^2 (\ln y_t - \ln \beta_0 - t \cdot \ln \beta_1) \cdot \frac{t}{\beta_1}$$

Anulováním pravých stran (56A), (56B) a po jednoduchých úpravách máme

$$\sum_{t=1}^n y_t^2 \ln y_t - \ln \beta_0 \sum_{t=1}^n y_t^2 - \ln \beta_1 \sum_{t=1}^n t y_t^2 = 0$$

$$\sum_{t=1}^n t \cdot y_t^2 \ln y_t - \ln \beta_0 \sum_{t=1}^n t \cdot y_t^2 - \ln \beta_1 \sum_{t=1}^n t^2 \cdot y_t^2 = 0$$

což po přemístění členů a vynásobení (-1) vede k (55A), (55B) . □ .

Soustavu (55A), (55B) lze maticově zapsat jako

$$\begin{pmatrix} \sum_{t=1}^n y_t^2 & \sum_{t=1}^n t y_t^2 \\ \sum_{t=1}^n t \cdot y_t^2 & \sum_{t=1}^n t^2 \cdot y_t^2 \end{pmatrix} \cdot \begin{pmatrix} \ln \beta_0 \\ \ln \beta_1 \end{pmatrix} = \begin{pmatrix} \sum_{t=1}^n y_t^2 \ln y_t \\ \sum_{t=1}^n t \cdot y_t^2 \ln y_t \end{pmatrix}$$

Odtud je snadno vidět, že má **explicitní řešení** pro logaritmované hodnoty obou parametrů ve tvaru

$$(56A) \quad \ln b_0 = \frac{\sum t^2 y_t^2 \cdot \sum y_t^2 \ln y_t - \sum t y_t^2 \cdot \sum t y_t^2 \ln y_t}{\sum t^2 y_t^2 \cdot \sum y_t^2 - (\sum t y_t^2)^2}$$

$$(56B) \quad \ln b_1 = \frac{\sum y_t^2 \cdot \sum t y_t^2 \ln y_t - \sum t y_t^2 \sum y_t^2 \ln y_t}{\sum t^2 y_t^2 \cdot \sum y_t^2 - (\sum t y_t^2)^2}$$

Odhady b_0, b_1 obou původních parametrů β_0, β_1 získáme exponenciálním povýšením výrazů na pravých stranách.

Vlastnost, že **exponenciální trend má koeficient růstu** a zároveň **podíl dvou sousedních diferencí** $(Tr_{t+2} - Tr_{t+1}) / (Tr_{t+1} - Tr_t)$ **konstantní**, plyne z následujícího:

Vyjdeme-li z (51) $Tr_t = \beta_0 \beta_1^t$, pak $y_{t+1} = \beta_0 \beta_1^{t+1}$ a tedy $Tr_{t+1} / Tr_t = \beta_1$ (konstantní)

a rovněž tak
$$\frac{Tr_{t+2} - Tr_{t+1}}{Tr_{t+1} - Tr_t} = \frac{\beta_0 \beta_1^{t+2} - \beta_0 \beta_1^{t+1}}{\beta_0 \beta_1^{t+1} - \beta_0 \beta_1^t} = \frac{\beta_0 \beta_1^{t+1} (\beta_1 - 1)}{\beta_0 \beta_1^t (\beta_1 - 1)} = \beta_1 .$$
 □ .

8A) výpočet parametrů modifikovaného exponenciálního trendu

$$(61) \quad Tr_t^{mex} = \beta_2 + \beta_0 \cdot \beta_1^t \quad [\beta_0 > 0] \text{ a zpravidla také } [0 < \beta_1 < 1].$$

Tento trend je tříparametrickým **zobecněním** předchozího případu (parametry $\beta_0, \beta_1, \beta_2$). Hodí se pro modelování **trendu s konstantním podílem sousedních diferencí**, pokud je navíc tento trend asymptoticky omezen (přibližuje-li se k saturační úrovni).

Jedna z orientačních metod odhadu parametrů **modifikovaného exponenciálního trendu** (pokud nemůžeme použít exaktní **iterační postupy nelineární optimalizace**), spočívá v tomto postupu:

Rozdělíme soubor pozorovaných hodnot na tři stejně velké třetiny o délce m . Pokud není n přesně dělitelné třemi, pak vynecháme jedno nebo dvě počáteční pozorování. Pak sečteme pozorování v jednotlivých třetinách, přičemž dostaneme

$$(62A) \quad \sum_1 y_t \approx \sum_1 T_t = m \cdot \beta_2 + \frac{\beta_0 \beta_1 (\beta_1^m - 1)}{\beta_1 - 1} \quad \sum_{t=1}^m y_t = m \beta_2 + \beta_0 \cdot \sum_{t=1}^m \beta_1^t$$

$$(62B) \quad \sum_2 y_t \approx \sum_2 T_t = m \cdot \beta_2 + \frac{\beta_0 \beta_1^{m+1} (\beta_1^m - 1)}{\beta_1 - 1} \quad \sum_{t=m+1}^{2m} y_t = m \beta_2 + \beta_0 \cdot \sum_{t=m+1}^{2m} \beta_1^t$$

$$(62C) \quad \sum_3 y_t \approx \sum_3 T_t = m \cdot \beta_2 + \frac{\beta_0 \beta_1^{2m+1} (\beta_1^m - 1)}{\beta_1 - 1} \quad \sum_{t=2m+1}^{3m} y_t = m \beta_2 + \beta_0 \cdot \sum_{t=2m+1}^{3m} \beta_1^t, \text{ kde}$$

indexování „1“ v $\sum_1 y_t$ resp. v $\sum_1 T_t$ značí součet pozorovaných hodnot, případně trendových hodnot z první třetiny časové řady. Řešením této **soustavy tří rovnic** dostaneme postupně jednotlivé odhady parametrů b_0, b_1, b_2 ve tvaru

$$(63A) \quad b_1 = \left(\frac{\sum_3 y_t - \sum_2 y_t}{\sum_2 y_t - \sum_1 y_t} \right)^{1/m}$$

$$(63B) \quad b_0 = \frac{b_1 - 1}{b_1 (b_1^m - 1)^2} (\sum_2 y_t - \sum_1 y_t)$$

$$(63C) \quad b_2 = \frac{1}{m} \left(\sum_1 y_t - \frac{b_0 b_1 (b_1^m - 1)}{b_1 - 1} \right)$$

Poznámka-odvození: při odhadu trojice parametrů postupujeme v tomto pořadí:

Pro určení parametru b_1 v (63A) nejprve odečteme rovnici (62A) od (62B):

$$\sum_3 y_t - \sum_2 y_t = m \cdot \beta_2 + \frac{\beta_0 \beta_1^{2m+1} (\beta_1^m - 1)}{\beta_1 - 1} - m \cdot \beta_2 - \frac{\beta_0 \beta_1^{m+1} (\beta_1^m - 1)}{\beta_1 - 1} = \frac{\beta_0 \beta_1^{m+1} (\beta_1^m - 1)^2}{\beta_1 - 1}$$

a následně odečteme rovnici (62B) od (62C):

$$\sum_2 y_t - \sum_1 y_t = m \cdot \beta_2 + \frac{\beta_0 \beta_1^{m+1} (\beta_1^m - 1)}{\beta_1 - 1} - m \cdot \beta_2 - \frac{\beta_0 \beta_1 (\beta_1^m - 1)}{\beta_1 - 1} = \frac{\beta_0 \beta_1 (\beta_1^m - 1)^2}{\beta_1 - 1}$$

Podílem obou vztahů dostaneme:

$$\frac{\sum_3 y_t - \sum_2 y_t}{\sum_2 y_t - \sum_1 y_t} = \frac{\beta_0 \beta_1^{m+1} (\beta_1^m - 1)^2 / (\beta_1 - 1)}{\beta_0 \beta_1 (\beta_1^m - 1)^2 / (\beta_1 - 1)} = \beta_1^m,$$

Nyní podobně vypočteme b_0 z rozdílu rovnic (62B) - (62A) při tentokrát již známém $b_1 = \beta_1$

$$\sum_2 y_t - \sum_1 y_t = m \cdot \beta_2 + \frac{\beta_0 b_1^{m+1} (b_1^m - 1)}{b_1 - 1} - m \cdot \beta_2 - \frac{\beta_0 b_1 (b_1^m - 1)}{b_1 - 1} = \frac{\beta_0 b_1 (b_1^m - 1)^2}{b_1 - 1},$$

takže určení b_0 bude snadno proveditelné pomocí výrazu (63B)

Parametr b_2 pak nakonec snadno získáme z (62A) při známých b_0, b_1 :

$$\sum_1 y_t \approx \sum_1 T_t = m \cdot \beta_2 + \frac{b_0 b_1 (b_1^m - 1)}{b_1 - 1}.$$

alternativní postup: Vzhledem k tomu, že při pevně zvoleném parametru b_1 se model stane **lineárním modelem**, lze také použít takový postup, že se odhadnou parametry b_0, b_2 pro různé hodnoty b_1 a následně se zvolí taková varianta, která minimalizuje SSE.

ověření indikace přítomnosti modifikovaného exponenciálního trendu:

podíly sousedních 1. diferencí $(y_{t+2} - y_{t+1}) / (y_{t+1} - y_t)$ **jsou zhruba konstantní:**

Protože zřejmě platí $y_{t+2} = \beta_2 + \beta_0 \cdot \beta_1^{t+2}$ $y_{t+1} = \beta_2 + \beta_0 \cdot \beta_1^{t+1}$ $y_t = \beta_2 + \beta_0 \cdot \beta_1^t$,

vyplývá odtud pro difference $y_{t+2} - y_{t+1} = \beta_0 \cdot \beta_1^{t+1} (\beta_1 - 1)$ $y_{t+1} - y_t = \beta_0 \cdot \beta_1^t (\beta_1 - 1)$ a následně $(y_{t+2} - y_{t+1}) / (y_{t+1} - y_t) = \beta_1$. Konstantnost představuje parametr β_1 .

Limitní chování modifikovaného exponenciálního trendu

Při přijatých omezeních na parametry funkce (61)

$$T_t = \beta_2 + \beta_0 \cdot \beta_1^t \quad [\beta_0 > 0], [\beta_2 > 0] \text{ a } [0 < \beta_1 < 1], \text{ platí}$$

$$\lim_{t \rightarrow 0} T_t = \lim_{t \rightarrow 0} (\beta_2 + \beta_0 \cdot \beta_1^t) = \beta_0 + \beta_2 \quad \lim_{t \rightarrow +\infty} T_t = \lim_{t \rightarrow +\infty} (\beta_2 + \beta_0 \cdot \beta_1^t) = \beta_2$$

Při přijatých omezeních na parametry funkce (61)

$$T_t = \beta_2 + \beta_0 \cdot \beta_1^t \quad [\beta_0 < 0] \text{ a } [0 < \beta_1 < 1], \text{ platí}$$

$$\lim_{t \rightarrow 0} T_t = \lim_{t \rightarrow 0} (\beta_2 + \beta_0 \cdot \beta_1^t) = \beta_0 + \beta_2 \quad \lim_{t \rightarrow +\infty} T_t = \lim_{t \rightarrow +\infty} (\beta_2 + \beta_0 \cdot \beta_1^t) = \beta_2$$

Chování **modifikovaného exponenciálního trendu** asymptoticky (pro $t \rightarrow +\infty$) směřuje k saturační úrovni dané hodnotou parametru β_2 , přičemž v blízkosti 0 by hodnoty mohly být i záporné (záleží na součtu parametrů $\beta_0 + \beta_2$).

9A) výpočet parametrů logistického trendu

Tento opět tříparametrický trend je popsán schématem

$$(71) \quad Tr_t^{lg} = \frac{\beta_2}{1 + \beta_0 \beta_1^t}, \quad t = 1, 2, 3, \dots, n \quad \text{při} \quad 1 > \beta_1 > 0, \beta_2 > 0$$

Logistický trend má *inflexi* (tj. bod, ve kterém přechází konvexní průběh trendu v konkávní, příp. naopak) v bodě $t^* = -\frac{\ln \beta_0}{\ln \beta_1}$ a je asymptoticky omezen (se saturační úrovní β_2).

Jeho zderivováním podle časové proměnné t (tu považujeme za spojitou), dostaneme ⁴

$$(72) \quad \frac{dTr_t^{lg}}{dt} = -\frac{\ln \beta_1}{\beta_2} Tr_t (\beta_2 - Tr_t), \quad \text{což je}$$

další důležitý ukazatel růstu trendové křivky (tato první derivace trendové křivky se v tomto kontextu nazývá *růstová funkce*). Je zřejmé, že rychlost růstu logistického trendu závisí přímo úměrně na dosažené úrovni Tr_t a na vzdálenosti hodnoty trendu od saturační úrovně β_2 , tj. $\beta_2 - Tr_t$. Derivace je přitom symetrická kolem bodu inflexe $t^* = -\ln \beta_0 / \ln \beta_1$. Z toho, co bylo řečeno, plyne, že logistickou křivku lze přiřadit k tzv. *s-křivkám symetrickým kolem inflexního bodu*.

Pokud jde o **odhad parametrů logistického trendu**, lze použití několik odlišných metod:

a) protože *logistický trend* lze považovat za *reciprokou podobu modifikovaného exponenciálního trendu*, lze aplikovat výše popsanou odhadovou proceduru pro odhad modifikovaného exp. trendu na časovou řadu s reciprokými hodnotami $1/y_t$.

b) Použijeme *diferenčního parametrického odhadu*, kdy se místo s původní časovou řadou pracuje s řadou prvních diferencí $y_{t+1} - y_t$. Přitom se postupuje takto:

V derivační formulaci trendu (72) se aproximuje trendová složka Tr_t skutečnými pozorováními y_t , takže se dostane:

$$(73) \quad \frac{dy_t}{dt} = -\frac{\ln \beta_1}{\beta_2} y_t (\beta_2 - y_t)$$

Jestliže dále přijmeme aproximaci (derivace pomocí diference)

$$(74) \quad \frac{dy_t}{dt} \approx \frac{y_{t+1} - y_t}{t + 1 - t} = y_{t+1} - y_t = \Delta y_t, \quad \text{v níž}$$

Δ_t označuje řadu prvních diferencí, pak ze (73) snadnou úpravou dostaneme

$$(75) \quad \frac{\Delta y_t}{y_t} = -\ln \beta_1 + \frac{\ln \beta_1}{\beta_2} \cdot y_t$$

$$\begin{aligned} \frac{\partial Tr_t^{lg}}{\partial t} &= \frac{\partial \beta_2 \cdot (1 + \beta_0 \beta_1^t)^{-1}}{\partial t} = -\beta_2 \frac{\beta_0 \beta_1^t \cdot \ln \beta_1}{(1 + \beta_0 \beta_1^t)^2} = -\frac{\beta_2 \ln \beta_1}{1 + \beta_0 \beta_1^t} \cdot \frac{\beta_0 \beta_1^t}{1 + \beta_0 \beta_1^t} = \\ &= -\frac{\beta_2 \ln \beta_1}{1 + \beta_0 \beta_1^t} \left(1 - \frac{1}{1 + \beta_0 \beta_1^t} \right) = -\frac{\ln \beta_1}{\beta_2} \frac{\beta_2}{1 + \beta_0 \beta_1^t} \left(\beta_2 - \frac{\beta_2}{1 + \beta_0 \beta_1^t} \right) = -\frac{\ln \beta_1}{\beta_2} Tr_t^{lg} (\beta_2 - Tr_t^{lg}) \end{aligned}$$

Pomocí **metody nejmenších čtverců** pak získáme v **lineárním regresním modelu** (modelu s jediným regresorem), tzn. ve stochastické formulaci (75)

$$(75^*) \quad \frac{\Delta y_t}{y_t} = -\ln \beta_1 + \frac{\ln \beta_1}{\beta_2} y_t + \varepsilon_t$$

$$(75^{**}) \quad \frac{y_{t+1} - y_t}{y_t} = \alpha_1 + \alpha_2 y_t + \varepsilon_t, \text{ kde } \alpha_1 = -\ln \beta_1 \quad \alpha_2 = \ln \beta_1 / \beta_2$$

odhady $\hat{\alpha}_1, \hat{\alpha}_2$ pro α_1, α_2 a odtud následně odhady parametrů β_1, β_2 jako

$$(76) \quad b_1 = \hat{\beta}_1 = \exp(-\hat{\alpha}_1) \text{ a pro } b_2 = \hat{\beta}_2 = \frac{-\hat{\alpha}_1}{\hat{\alpha}_2}$$

Abychom získali odhad pro poslední parametr β_0 , aproximujeme ve vztahu (71) trendovou složku Tr_t skutečnými pozorováními y_t a upravíme tento vztah do tvaru

$$(77) \quad \beta_0 \beta_1^t \approx \frac{\beta_2}{y_t} - 1 \quad (\text{protože } Tr_t \approx y_t = \frac{\beta_2}{1 + \beta_0 \beta_1^t})$$

Po zlogaritmování $\ln \beta_0 + t \ln \beta_1 = \ln \left(\frac{\beta_2}{y_t} - 1 \right)$ a po sečtení těchto vztahů přes $t = 1, 2, \dots, n$ ⁵

dostaneme nakonec vztah

$$(78) \quad \ln \beta_0 = -\frac{(n+1) \ln \beta_1}{2} + \frac{1}{n} \sum_{t=1}^n \ln \left(\frac{\beta_2}{y_t} - 1 \right)$$

který se nazývá **Rhodesův vzorec**.

Z něho se už snadno vypočte odhad parametru β_0 „vyexponováním“.

Ověření toho, že **logistický trend** je **reciproké vyjádření modifikovaného exponenciálního trendu** je velmi jednoduché:

stačí (71) $Tr_t^{lg} = \frac{\beta^{*2}}{1 + \beta_0^* \beta_1^{*t}}$ zapsat v reciproké podobě $\frac{1}{Tr_t^{lg}} = \frac{1 + \beta_0^* \beta_1^{*t}}{\beta^{*2}}$ a hned vidíme,

že výraz odpovídá zápisu modifikovaného exponenciálního trendu (61) $Tr_t^{mex} = \beta_2 + \beta_0 \cdot \beta_1^t$,

pokud vyjádříme parametry tohoto trendu jako $\beta_2 = \frac{1}{\beta_2^*}, \beta_0 = \frac{\beta_0^*}{\beta_2^*}, \beta_1 = \beta_1^*$

⁵ Z $\sum_{t=1}^n (\ln \beta_0 + t \ln \beta_1) = \sum_{t=1}^n \ln \left(\frac{\beta_2}{y_t} - 1 \right)$ dostaneme $n \ln \beta_0 + \ln \beta_1 \frac{n(n+1)}{2} = \sum_{t=1}^n \ln \left(\frac{\beta_2}{y_t} - 1 \right)$ a dělíme n

Ověření lokalizace inflexního bodu u logistické křivky

Vzhledem k tomu, že inflexe je místem, kde konvexnost přechází v konkávnost (nebo vice versa) musíme spočítat druhou derivaci (podle času t). trendové křivky. Postupně dostaneme:

$$(79A) \quad \frac{dy_t}{dt} = -\beta_2 \cdot (1 + \beta_0 \beta_1^t)^{-2} \cdot \beta_0 \beta_1^t \ln \beta_1 = -\beta_2 \frac{\beta_0 \beta_1^t \cdot \ln \beta_1}{(1 + \beta_0 \beta_1^t)^2} \quad \text{a následně}$$

$$\frac{d^2 y_t}{dt^2} = -\beta_2 \cdot \frac{(1 + \beta_0 \beta_1^t)^2 \cdot [\beta_0 \beta_1^t \ln^2 \beta_1] - [\beta_0 \beta_1^t \ln \beta_1]^2 \cdot 2(1 + \beta_0 \beta_1^t)}{(1 + \beta_0 \beta_1^t)^4} \quad \text{neboli}$$

$$(79B) \quad \frac{d^2 y_t}{dt^2} = -\beta_2 \cdot (1 + \beta_0 \beta_1^t) \left[\beta_0 \beta_1^t \ln \beta_1 \right] \frac{\ln \beta_1 - \beta_0 \beta_1^t \ln \beta_1}{(1 + \beta_0 \beta_1^t)^4}$$

Nulovou hodnotu může druhá derivace nabýt jen tam, kde je obsah čitatele zlomku nulový, protože při daných omezeních na parametry $\beta_0 > 0, \beta_1 > 0$ nemůže pro žádné reálné t platit

$1 + \beta_0 \beta_1^t = 0$. Pro nulový čítecil vyplývá tedy navazující podmínka

$$\frac{d^2 y_t}{dt^2} = 0 \Rightarrow \ln \beta_1 - \beta_0 \beta_1^t \ln \beta_1 = 0, \text{ resp. } \ln \beta_1 (1 - \beta_0 \beta_1^t) = 0$$

a vzhledem k nenulovosti $\ln \beta_1$ pro $\beta_1 \neq 1$ může mít tento součin hodnotu 0 jen tam, kde platí

$1 - \beta_0 \cdot \beta_1^t = 0$ neboli $\beta_0 \cdot \beta_1^t = 1$ a tedy $\beta_1^t = \frac{1}{\beta_0}$. Hodnotu t^* lokalizující inflexi pak už

snadno získáme zlogaritmováním: $t \cdot \ln \beta_1 = \ln 1 - \ln \beta_0 = -\ln \beta_0$, odkud odvodíme polohu

inflexního bodu $t^* = \frac{-\ln \beta_0}{\ln \beta_1}$ □ .

ověření indikace přítomnosti logistického trendu:

podíly sousedních 1. diferencí recipročných hodnot $(1/y_{t+2} - 1/y_{t+1}) / (1/y_{t+1} - 1/y_t)$ jsou zhruba konstantní.

Protože platí $\frac{1}{y_{t+2}} = \frac{1 + \beta_0 \beta_1^{t+2}}{\beta_2}$, $\frac{1}{y_{t+1}} = \frac{1 + \beta_0 \beta_1^{t+1}}{\beta_2}$, $\frac{1}{y_t} = \frac{1 + \beta_0 \beta_1^t}{\beta_2}$ vyplývá odtud pro

Diference $y_{t+2}^{-1} - y_{t+1}^{-1} = \beta_2^{-1} \beta_0 \cdot \beta_1^{t+1} (\beta_1 - 1)$, $y_{t+1}^{-1} - y_t^{-1} = \beta_2^{-1} \beta_0 \cdot \beta_1^t (\beta_1 - 1)$ a

tedy opravdu platí $(1/y_{t+2} - 1/y_{t+1}) / (1/y_{t+1} - 1/y_t) = \beta_1$ □ .

Konstantnost tedy představuje stejně jako u modifikovaného exponenciálního trendu parametr β_1 .

10A) Gompertzův trend a výpočet jeho parametrů

Trend ve tvaru této křivky vzniká (ostatně stejně jako logistický trend) transformací **modifikovaného exponenciálního trendu**.

$$(81) \quad \ln Tr_t = \beta_3 + \beta_1 \cdot \beta_2^t \quad \text{resp. ekvivalentně} \quad Tr_t = \exp(\beta_3 + \beta_1 \cdot \beta_2^t)$$

při $[\beta_2 > 0] \quad t = 1, 2, \dots, n$

Obvykle – pro dosažení charakteristického tvaru křivky – **přijímáme**: $\beta_1 < -1, 0 < \beta_2 < 1$

Pro hodnoty parametrů z tohoto intervalu má **Gompertzův trend** inflexi v bodě

$$(82) \quad t^* = -\frac{\ln(-\beta_1)}{\ln(\beta_2)}$$

První derivace této křivky (tzv. **růstová funkce**) ale tentokrát není symetrická kolem inflexního bodu, ale **je sešikmená** (protáhlejší) **doprava**. Z tohoto důvodu se Gompertzův trend řadí mezi tzv. **s-křivky nesymetrické kolem inflexního bodu**.

Odhad parametrů $\beta_1, \beta_2, \beta_3$ se provede způsobem obdobným jako v případě modifikovaného exponenciálního trendu, místo původní řady se ale použije logaritmovaná původní řada (s hodnotami $\ln y_t$ ⁶)

Ověření lokalizace inflexního bodu u Gompertzovy křivky

Vzhledem k tomu, že inflexe je místem, kde konvexnost přechází v konkávnost (nebo vice versa) musíme spočítat druhou derivaci (podle času t). trendové křivky. Postupně dostaneme:

$$(83) \quad \frac{dy_t}{dt} = \exp(\beta_3 + \beta_1 \cdot \beta_2^t) \cdot \frac{d(\beta_3 + \beta_1 \cdot \beta_2^t)}{dt} = \exp(\beta_3 + \beta_1 \cdot \beta_2^t) \cdot \beta_1 \cdot \beta_2^t \cdot \ln \beta_2$$

a následně
$$\frac{d^2 y_t}{dt^2} = \exp(\beta_3 + \beta_1 \cdot \beta_2^t) \left(\beta_1 \cdot \beta_2^t \cdot \ln \beta_2 \right)^2 + \exp(\beta_3 + \beta_1 \cdot \beta_2^t) \cdot \beta_1 \cdot \beta_2^t \cdot \ln^2(\beta_2)$$

Vytkneme-li z obou členů (nenulový) výraz s exponenciálou, obdržíme

$$(84) \quad \frac{d^2 y_t}{dt^2} = \exp(\beta_3 + \beta_1 \cdot \beta_2^t) \left[\left(\beta_1 \cdot \beta_2^t \cdot \ln \beta_2 \right)^2 + \beta_1 \cdot \beta_2^t \cdot \ln^2 \beta_2 \right]$$

Nulovou hodnotu může druhá derivace nabýt jen tam, kde je obsah hranaté závorky nulový, tedy

v bodě splňujícím podmínku $\frac{d^2 y_t}{dt^2} = 0 \Rightarrow \beta_1^2 \cdot \beta_2^{2t} \cdot \ln^2 \beta_2 + \beta_1 \cdot \beta_2^t \cdot \ln^2 \beta_2 = 0$, po vytknutí

tedy $\ln^2(\beta_2) \beta_1 \beta_2^t [\beta_1 \cdot \beta_2^t + 1] = 0$ přičemž vzhledem k nenulovosti součinu před závorkou

může tento výraz nabýt nulovou hodnotu je tam, kde je obsah hranaté závorky nulový, tj. tam, co

platí $\beta_1 \cdot \beta_2^t + 1 = 0$ neboli $\beta_1 \cdot \beta_2^t = -1$ a tedy $\beta_2^t = \frac{1}{-\beta_1}$. Hodnotu t^* lokalizující inflexi pak

už snadno získáme zlogaritmováním:

$$t \cdot \ln \beta_2 = \ln(1) - \ln(-\beta_1) = -\ln(-\beta_1), \text{ odkud máme polohu inflexe } t^* = \frac{-\ln(-\beta_1)}{\ln \beta_2} \quad \square .$$

⁶ Je tímto zřejmé, že postup nelze uplatnit na ekonomickou časovou řadu s některými hodnotami zápornými.

ověření indikace přítomnosti Gompertzova trendu:

indikace přítomnosti: podíly prvních logaritmovaných diferencí $(\ln y_{t+2} - \ln y_{t+1}) / (\ln y_{t+1} - \ln y_t)$

jsou přibližně konstantní

Z **logaritmického vyjádření Gompertzova trendu (81)** dostaneme pro $t+1, t+2$

$$\ln y_{t+1} = \beta_3 + \beta_1 \cdot \beta_2^{t+1} \quad \ln y_{t+2} = \beta_3 + \beta_1 \cdot \beta_2^{t+2}, \text{ a následně tedy pro první}$$

dvě logaritmované diference:

$$\ln y_{t+2} - \ln y_{t+1} = \beta_1 \cdot \beta_2^{t+2} - \beta_1 \cdot \beta_2^{t+1} = \beta_1 \cdot \beta_2^{t+1} (\beta_2 - 1)$$

$$\ln y_{t+1} - \ln y_t = \beta_1 \cdot \beta_2^{t+1} - \beta_1 \cdot \beta_2^t = \beta_1 \cdot \beta_2^t (\beta_2 - 1). \text{ Odtud máme pro jejich podíl}$$

$$\frac{\ln y_{t+2} - \ln y_{t+1}}{\ln y_{t+1} - \ln y_t} = \frac{\beta_1 \cdot \beta_2^{t+1} (\beta_2 - 1)}{\beta_1 \cdot \beta_2^t (\beta_2 - 1)} = \beta_2, \text{ takže ona přibližná konstanta je parametr } \beta_2.$$

1J) výpočet parametrů odmocninného trendu $y_t = \beta_0 + \beta_1 \cdot \sqrt{t}$

Aplikuje-li standardní **OLS – odhadové schéma** pro tento model, který je lineární v parametrech, dostaneme

$$(91) \quad \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} n & \sum_{t=1}^n \sqrt{t} \\ \sum_{t=1}^n \sqrt{t} & \sum_{t=1}^n t \end{pmatrix}^{-1} \begin{pmatrix} \sum_{t=1}^n y_t \\ \sum_{t=1}^n \sqrt{t} \cdot y_t \end{pmatrix}$$

$$\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} n & \sum_{t=1}^n \sqrt{t} \\ \sum_{t=1}^n \sqrt{t} & \sum_{t=1}^n t \end{pmatrix}^{-1} \begin{pmatrix} \sum_{t=1}^n y_t \\ \sum_{t=1}^n y_t \sqrt{t} \end{pmatrix} = \frac{1}{\frac{n^2(n+1)}{2} - \left(\sum_{t=1}^n \sqrt{t}\right)^2} \begin{pmatrix} \frac{n(n+1)}{2} & -\sum_{t=1}^n \sqrt{t} \\ -\sum_{t=1}^n \sqrt{t} & n \end{pmatrix} \begin{pmatrix} \sum_{t=1}^n y_t \\ \sum_{t=1}^n y_t \sqrt{t} \end{pmatrix},$$

takže pro parametrické odhady (budou konzistentní a nestranné) máme

$$b_0 = \hat{\beta}_0 = \frac{\frac{n(n+1)}{2} \left(\sum_{t=1}^n y_t\right) - \left(\sum_{t=1}^n \sqrt{t}\right) \left(\sum_{t=1}^n y_t \sqrt{t}\right)}{\frac{n^2(n+1)}{2} - \left(\sum_{t=1}^n \sqrt{t}\right)^2}, \quad b_1 = \hat{\beta}_1 = \frac{n \left(\sum_{t=1}^n y_t \sqrt{t}\right) - \left(\sum_{t=1}^n \sqrt{t}\right) \left(\sum_{t=1}^n y_t\right)}{\frac{n^2(n+1)}{2} - \left(\sum_{t=1}^n \sqrt{t}\right)^2},$$

V případech, kdy je levostranná regresní proměnná transformací původní veličiny časové řady y_t , lze aplikovat metodu nejmenších čtverců OLS jen s výhradou.

Minimalizace reziduí prováděná pomocí OLS nebude totiž založena na kritériu Mine_t^2 , nýbrž na kritériu $\text{Min} [f(e_t)]^2$, kde $f(\cdot)$ je použitá transformující funkce, nejčastěji logaritmická funkce nebo exponenciála. Pro takovéto případy by měla být užita nelineární (prostá) metoda nejmenších čtverců **NLLS (Non-Linear Least Squares Method)**, jejímž minimalizačním kritériem je právě

$$\text{Min} \sum_{t=1}^n [f(y_t) - f(\hat{y}_t)]^2$$

Výpočet je ale nutné provést některou z **numerických iteračních metod**, protože výsledný vzorec pro odhadnuté parametry nelze vyjádřit explicitně.

Hlavní zásady výběru výstižného trendu

- 1) Zpravidla preferujeme trendovou křivku s co nejmenším počtem parametrů (obvykle do 3). Zásadou je úspornost (*parsimony*), která je zde ještě více opodstatněná než v klasické regresní analýze, neboť vodítka pro tvar nelinearity zde až na ojedinělé výjimky nemůžeme vyvodit z ekonomické teorie. Jednou z možností, jak otestovat potřebu/nepotřebu nutnosti zařazení další trendové transformace může být **výpočet korigovaného koeficientu determinace \tilde{R}^2** . Jinou racionální úvahu lze založit na určení statistické významnosti příslušného trendového parametru (k tomu lze užít klasické testování pomocí t-testu). V trendové analýze není zdaleka tak osudové riziko specifikační chyby spočívající ve vynechání některé z relevantních proměnných jako je tomu v analýze regresní.
- 2) Před první volbou trendové křivky je vhodné se podívat na graf pozorovaných hodnot časové řady, který bývá prvotním vodítkem výchozí trendové specifikace. Z něho lze dost často vyvodit řád nestacionarity např. tím, že vytvoříme řadu prvních nebo druhých diferencí. Pokud ani diference vyšších řádů nevedou ke stacionární řadě a řada vykazuje monotónní akcelerující růst, je to indikací pro volbu *exponenciálního trendu* (s případným posunem výchozí úrovně). Určitým indikátorem je pak graf průběhu reziduálních hodnot, z něhož lze často vyvodit alší transformující funkci času, která byla v základní specifikaci opomenuta (např. mnohočlen 2. nebo 3 stupně).
- 3) Nesmíme zapomenout na to, že **trendová analýza slouží především ke krátkodobým predikcím**. Vzhledem k tomu, že polynomy 4. a vyšších řádů se vyznačují značnou citlivostí při predikcích, je to důvodem pro jejich vynechání z oboru možných trendových funkcí. (Ekonomické ukazatele se chovají principiálně ustáleněji než např. fyzikální/technické, kde může být účelné nasazení např. *splínových funkcí*).
- 4) Pokud ekonomický ukazatel vykazuje evidentně rozdílný průběh ve sledovaném minulém období, obvykle to velmi komplikuje rozhodnutí o výběru trendu. V takovémto případě – pokud se nevzdáme pokusu o predikci vůbec – přistupujeme zpravidla k prolongování chování z údajů posledního období nebo z údajů toho minulého období, o němž lze důvodně předpokládat, že vývoj zde pozorovaný se vyskytne i v nejbližší budoucnosti.
- 5) Přítomnost *heteroskedasticity* (standardně testované v regresní analýze) není v případě trendové analýzy časové řady zpravidla vážnějším problémem.⁷ U testování konvenčními testy nevzniká specifický problém, jen u *Whiteova testu* může být větší než obvykle pravděpodobnost výskytu součinné kombinace dvou elementárních trendových funkcí s některou z dalších zařazených (vznikla by dokonalá multikolinearita). U *Goldfeld-Quandtova testu* se zřejmě nemůžeme opřít o obsahovou vazbu variability reziduí s proměnlivostí některé z vysvětlujících proměnných, takže jeho nasazení ztrácí původní význam. V testování převažují prvky technické nad věcnými.

⁷ Pracujeme s časovou řadou, jejíž hodnoty – aspoň v nedlouhém časovém období – jsou zpravidla dosti vzájemně blízké a rozptýl náhodných složek obvykle významněji nekolísá.

6) Naopak, důležitým prvkem trendové analýzy je vyšetřování síly a stupně *autokorelace náhodných složek* sledované časové řady. Náhodné složce časové řady je vůbec věnována značná pozornost. Proto také obvykle první pohled na výsledky směřuje k vyšetření reziduálních hodnot, přičemž základní představu o jejich případné autokorelovanosti poskytne Durbin-Watsonův koeficient autokorelace 1.řádu.

7) Další z průvodních problémů standardní lineární regrese – *multikolinearita* – se v trendové analýze vyskytnout nemůže: K *dokonalé multikolinearitě* nemůže vést žádná lineární kombinace elementárních trendových komponent (jsou-li tyto rozdílné) a eventualita výskytu *přibližné multikolinearity* nemůže pocházet ze *stochastičnosti* „vysvětlujících proměnných,, (jsou-li elementární trendové komponenty svou povahou nestochastické).

8) Případný přínos apriorní informace: Je poměrně skromný počet případů, kdy lze k výběru nejvhodnější trendové funkce uplatnit poznatky plynoucí z externí ekonomické reality. Přece však některé existují: Ve vývoji průměrné mzdy jako nejcharakterističtější případ lze pozorovat tendenci pravidelného každoročního procentního nárůstu (který ovšem nemusí být každoročně stejný). Je to přímým důsledkem „zvyklostí“, které jsou zažité při mzdovém vyjednávání odborářů (a od nichž se sekundárně odvíjí i mzdový nárůst pracovníků ve státních podnicích a státní správě). **Obdobně, pokud by valorizace důchodů probíhala tak, že bude každoročně přidávána konstantní hodnota v Kč, bylo by to indikací lineárního trendu. Pokud by se přidávalo „o zhruba stejné procento“ (jako je to zvykem u mezd) , signalizovalo by to průběh vyjádřený exponenciální křivkou.**

Kritéria posuzování přesnosti předpovědi

Střední chyba odhadu [mean error] ME

$$ME = \frac{\sum_{t=n+1}^{n+h} (y_t - \hat{y}_t)}{h}$$

$$ME = \frac{\sum_{t=1}^n (y_t - \hat{T}r_t)}{n}$$

Jde o obvykle velmi malou hodnotu (řádů 10^{-8} a níže) indikující „míru nelinearity“ při výpočtu. Pro posouzení vlastní míry přesnosti predikční výpovědi žádný smysl nemá.

Součet čtvercových chyb odhadu [sum of squared errors] SSE

$$SSE = \sum_{t=n+1}^{n+h} (y_t - \hat{y}_t)^2 = \sum_{t=n+1}^{n+h} e_t^2$$

$$SSE = \sum_{t=1}^n (y_t - \hat{T}r_t)^2 = e_t^2$$

Je to stejné kritérium, jako kritérium nejmenších čtverců v regresním modelu (lineárním i nelineárním). Lze ho rozumně použít jen při srovnání uvažované s jinak specifikovanou trendovou funkcí (pro shodný počet pozorování). Měrovou jednotkou této míry je zde „čtverec“, původní měrné jednotky.

Střední čtvercová chyba odhadu [mean squared error] MSE

$$MSE = \frac{1}{h} \sum_{t=n+1}^{n+h} (y_t - \hat{y}_t)^2 = \frac{1}{h} \sum_{t=n+1}^{n+h} e_t^2$$

$$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{T}r_t)^2$$

se uplatňuje často coby **kvadratická ztrátová** („loss,“) *funkce*. Lze ji dekomponovat na tři složky takto:

$$MSE = \frac{1}{h} \sum_{t=n+1}^{n+h} (y_t - \hat{y}_t)^2 = (\bar{\hat{y}} - \bar{y})^2 + (s_{\hat{y}} - s_y)^2 + 2(1 - r_{y\hat{y}})s_{\hat{y}}s_y, \text{ kde na pravé straně je}$$

proporční vychýlení

$$\frac{(\bar{\hat{y}} - \bar{y})^2}{\frac{1}{h} \sum_{t=n+1}^{n+h} (y_t - \hat{y}_t)^2}$$

vzdálenost průměru předpovědí od průměru

předpovídáných hodnot

proporční rozptyl

$$\frac{(s_{\hat{y}} - s_y)^2}{\frac{1}{h} \sum_{t=n+1}^{n+h} (y_t - \hat{y}_t)^2}$$

vzdálenost rozptylu předpovědí od rozptylu

předpovídáných hodnot

proporční kovariance

$$\frac{2(1 - r_{y\hat{y}})s_{\hat{y}}s_y}{\frac{1}{h} \sum_{t=n+1}^{n+h} (y_t - \hat{y}_t)^2}$$

pokrývá zbývající nesystematickou část předpovědní chyby.

U dobré předpovědní techniky jsou první dvě složky relativně malé a výrazně převládá nesystematická složka. Součet všech tří z jich je 1. Odvození se získá z rozkladu

$$\frac{1}{h} \sum_{t=n+1}^{n+h} (y_t - \hat{y}_t)^2 = \frac{1}{h} \sum_{t=n+1}^{n+h} (y_t - \bar{y} + \bar{y} + \hat{y}_t - \bar{\hat{y}} + \bar{\hat{y}})^2 = \frac{1}{h} \sum_{t=n+1}^{n+h} (\bar{\hat{y}} - \bar{y} + \hat{y}_t - \bar{\hat{y}} + y_t - \bar{y})^2$$

Odmocninová střední čtvercová chyba [root mean squared error] RMSE

$$RMSE = \sqrt{\frac{1}{h} \sum_{t=n+1}^{n+h} (y_t - \hat{y}_t)^2} = \sqrt{\frac{1}{h} \sum_{t=n+1}^{n+h} e_t^2}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{T}r_t)^2}$$

Výhodou oproti **MSE** je vyjádření ve stejných jednotkách jako mají původní data.

Střední absolutní chyba odhadu [mean absolute error] MAE

$$MAE = \frac{\sum_{t=n+1}^{n+h} |y_t - \hat{y}_t|}{h}$$

$$MAE = \frac{\sum_{t=1}^n |y_t - \hat{T}r_t|}{n}$$

Její předností je – zejména v datových souborech s výraznými odchylkami některých pozorování od „standardních“ (tzv. **outliers**) - že nepenalizuje tak silně jako **MSE** či **RMSE** tyto „extrémní“ odchylky.

Dále uváděné míry již na měřítku hodnot předpovídané proměnné nezávisí:

Střední absolutní procentuální chyba odhadu [mean absolute percentage error] MPE

$$MAPE = \frac{100}{h} \cdot \sum_{t=n+1}^{n+h} \frac{|y_t - \hat{y}_t|}{y_t}$$

$$MAPE = \frac{100}{n} \cdot \sum_{t=1}^n \frac{|y_t - \hat{T}r_t|}{y_t}$$

Zpravidla nabývá hodnot mezi 0-100%, přičemž výsledek menší/lepší než 100% říká, že daný předpovědní model je lepší než model náhodné procházky s předpověďmi trvale na nulové úrovni, tj. trvale s MAPE=100%. Kritérium je však problematicky použitelné, pokud hodnoty časové řady jsou příliš malé (oscilují kolem 0).

Korigovaná střední absolutní procentuální chyba odhadu [adjusted mean absolute percentage error] AMAPE

$$AMAPE = \frac{100}{h} \cdot \sum_{t=n+1}^{n+h} \left| \frac{y_t - \hat{y}_t}{(y_t + \hat{y}_t)/2} \right|$$

$$AMAPE = \frac{100}{h} \cdot \sum_{t=1}^n \left| \frac{y_t - \hat{T}r_t}{(y_t + \hat{T}r_t)/2} \right|$$

Koriguje asymetrii předchozího kritéria **MAPE**: dává totiž stejný výsledek i při prohození skutečných a předpovídaných hodnot v tomto smyslu: Skutečná hodnota 0,6 a předpověď 0,8 přispějí do součtu v **AMAPE** stejně jako skutečná hodnota 0,8 a předpověď 0,6.

Střední procentní chyba odhadu [mean percentage error] MPE

$$MPE = \frac{100}{n} \cdot \sum_{t=n+1}^{n+h} \left(\frac{y_t - \hat{y}_t}{y_t} \right)$$

$$MPE = \frac{100}{n} \cdot \sum_{t=1}^n \left(\frac{y_t - \hat{T}r_t}{y_t} \right)$$

Je víceméně doplňkovým kritériem k **ME**, kdy v součtu vystupují relativní odchylky místo absolutních.

Theilův koeficient nesouladu [Theil's U-coefficient] ThU⁸

$$ThU = \frac{\sqrt{\sum_{t=n+1}^{n+h} (y_t - \hat{y}_t)^2}}{\sqrt{\sum_{t=n+1}^{n+h} y_t^2} + \sqrt{\sum_{t=n+1}^{n+h} \hat{y}_t^2}}, \text{ kde}$$

⁸ Henri Theil byl významný nizozemský ekonometr a matematický ekonom [1924-2000], následník Jana Tinbergena na Erasmus University v Rotterdamu, později působil na universitě v Chicagu a na University of Florida.

n délka původní časové řady (užité pro odhad modelu)

h délka předpovídaného období [$h > 0$]

$n + h$ délka původní časové řady prodloužená o předpovídané období.

Někdy je ovšem toto kritérium formulováno obecněji, přičemž odchylky jsou brány vůči nějakému jednoduchému či naivnímu modelu- tzv. **benchmark**. Hodnota U leží vždy mezi 0 a 1, přičemž hodnota 0 značí perfektní shodu předpovědi se skutečností.

Publikace Hindls R., Hronová S, Novák, I.: **Metody statistické analýzy pro ekonomy** uvádí jinou verzi **Theilova koeficientu** (zde označenou **ThC**)

Theilův koeficient nesouladu [Theil's C-coefficient] ThC

$$ThC = \frac{\sum_{t=n+1}^{n+h} (y_t - \hat{P}_t)^2}{\sum_{t=n+1}^{n+h} y_t^2}, \text{ kde}$$

n délka původní časové řady (užité pro odhad modelu)

h délka časové řady po zkrácení [$h > 0$] o q pozorovaných hodnot

$n + h$ délka původní časové řady prodloužená o předpovídané období.

\hat{P}_i extrapolovaná hodnota na i období dopředu, a to modelem odhadnutým na základě prvních r pozorování časové řady .

Pokud se hodnota ThC nachází v rozmezí 3-5%, lze mluvit o velmi uspokojivém výsledku. Model lze s velkou pravděpodobností považovat za použitelný k předpovědím.

Pokud se hodnota ThC nachází v rozmezí 5-10%, lze mluvit o jakžtakž uspokojivém výsledku. Model lze s mírnou nadějí považovat za použitelný k předpovědím

Pokud se hodnota ThC nachází nad úroveň 10%, nelze mluvit o uspokojivém výsledku. Model lze téměř s jistotou zavrhnout z hlediska použitelnosti k předpovědím.

K předchozím kritériím lze přidat ještě dvě další míry, u nichž se omezujeme na posouzení toho, jak model předpovídá znaménka budoucích hodnot (tj. zda tyto hodnoty budou kladné nebo záporné) nebo změny ve směru vývoje budoucích hodnot (zda růst přejde v pokles apod.) Někdy – při strategických úvahách – jde o důležitější aspekty než číselné co nejpřesnější vyjádření předpovídaných hodnot.

Procento správných předpovědí znaménka PCPS

$$PCPS = \frac{100}{h} \cdot \sum_{t=n+1}^{n+h} z_i, \text{ kde } z_i = 1 \text{ pro } y_t \cdot \hat{y}_t > 0$$

$$z_i = 0 \quad \text{pro} \quad y_t \cdot \hat{y}_t \leq 0$$

Charakteristika (zřejmě v rozmezí 0% -100%) není nijak moc vypovídající, neboť pouze informuje o tom, zda mají skutečné a predikované hodnoty stejná znaménka. V případě, že se hodnoty analyzované časové řady pohybují ve vysokých kladných číslech (běžná situace u velké části ekonomických ukazatelů) je prakticky bezcenná (hodnota **PCPS** bude s velkou pravděpodobností identicky 100).

Procento správných (jednokrokových) předpovědí změn směru vývoje PCPD

$$PCPD = \frac{100}{h} \cdot \sum_{t=n+1}^{n+h} z_i, \text{ kde } z_i = 1 \quad \text{pro} \quad (y_t - y_{t-1})(\hat{y}_t - y_{t-1}) > 0$$

$$z_i = 0 \quad \text{pro} \quad (y_t - y_{t-1})(\hat{y}_t - y_{t-1}) \leq 0$$

Na rozdíl od předchozího **PCPS** respektuje tento indikátor podstatně více charakter posuzování předpovědí u ekonomických časových řad. Posuzuje poměr stejnosměrnosti vůči protisměrnosti skutečného vývoje a předpovědi (oproti y_{t-1}) v situaci, kdy známe skutečnou hodnotu řady v minulém období $t - 1$. Vystižení správného směru vývoje z hodnoty y_{t-1} predikcí (růst /pokles) je někdy důležitější než samotné (přesné) určení konkrétní hodnoty.

Trendy s konstantními parametry

Trendem bývá obvykle nazýván **relativně ustálený a déleodobý vývoj časové řady ekonomického ukazatele/procesu oproštěný od dalších působících vlivů**

(sezónnosti, náhodnosti, eventuálně cykličnosti), v němž se projevují **systematické tendence chování jevu/jevů**, který daný ukazatel zobrazuje.

Je obvykle v podstatných rysech vystižitelný matematickým schématem. Na rozdíl od regresního schématu se však zde nezabýváme souvislostí chování sledovaného ukazatele s jinými proměnnými (ekonomickými i jinými), nýbrž posuzujeme vývoj ukazatele víceméně izolovaně. „Vysvětlující/mi proměnnou/nými je/ jsou zde pouze časové plynutí, případně upravené elementárními matematickými transformacemi časové proměnné t .

Mluvíme-li o **konstantních parametrech trendu**, rozumíme tím, že ve sledovaném časovém nedochází k podstatnějším „strukturálním“ změnám (za ty nepokládáme nahodilé odchylky), které by znamenaly zásadnější proměnu globální vývojové tendence resp. výraznou změnu chování analyticky vyjádřené matematické křivky. Trend může mít i časová řada, jejíž převažující vývojová tendence není vyjádřitelná

analytickým zápisem, případně i ta, kdy tato tendence není popsatelná

matematicky formulovatelnou závislostí vůbec. Ve druhém případě se obvykle

uchylujeme k vyjádření pomocí přibližného formálního schématu.

Určení trendu má smysl jak pro poznání dosavadního chování časové řady, tak a zejména pro – krátkodobou, případně až střednědobou – předpověď jejího dalšího vývoje.

Limitní chování modifikovaného exponenciálního trendu

Při přijatých omezeních na parametry funkce (61)

$$T_t = \beta_2 + \beta_0 \cdot \beta_1^t \quad [\beta_0 > 0], [\beta_2 > 0] \text{ a } [0 < \beta_1 < 1], \text{ platí}$$

$$\lim_{t \rightarrow 0} T_t = \lim_{t \rightarrow 0} (\beta_2 + \beta_0 \cdot \beta_1^t) = \beta_0 + \beta_2 \quad \lim_{t \rightarrow +\infty} T_t = \lim_{t \rightarrow +\infty} (\beta_2 + \beta_0 \cdot \beta_1^t) = \beta_2$$

Při přijatých omezeních na parametry funkce (61)

$$T_t = \beta_2 + \beta_0 \cdot \beta_1^t \quad [\beta_0 < 0] \text{ a } [0 < \beta_1 < 1], \text{ platí}$$

$$\lim_{t \rightarrow 0} T_t = \lim_{t \rightarrow 0} (\beta_2 + \beta_0 \cdot \beta_1^t) = \beta_0 + \beta_2 \quad \lim_{t \rightarrow +\infty} T_t = \lim_{t \rightarrow +\infty} (\beta_2 + \beta_0 \cdot \beta_1^t) = \beta_2$$

Chování modifikovaného exponenciálního trendu asymptoticky (pro $t \rightarrow +\infty$) směřuje k saturační úrovni dané hodnotou parametru β_2 , přičemž v blízkosti 0 by hodnoty mohly být i záporné (záleží na součtu parametrů $\beta_0 + \beta_2$).