Modely s rozloženými zpožděními I

Modely s rozloženými zpožděními (též „**modely rozdělených zpoždění**") tvoří důležitou (a dosti rozsáhlou) oblast regresních schémat využívaných v ekonometrii. Všechny se vyznačují **přítomnost**í **zpožděných proměnných** (a to jak **vysvětlujících** proměnných, tak někdy **i vysvětlované** proměnné) mezi ostatními (nezpožděnými) veličinami regresní rovnice. Tyto modely současně znamenají začlenění dynamických prvků do modelů analyzovaných ekonometrickými metodami.

Lze je klasifikovat několika způsoby, jedno z možných rozdělení je toto :
A. modely s konečnou délkou zpoždění (modely s konečným zpožděním)
B. modely se zpožděním o nekonečné délce (modely s nekonečným zpožděním)

S uplatněním modelů rozložených zpoždění se pojí dvě základní otázky :

a) **jaká maximální délka (hloubka) zpoždění je ještě únosná pro zařazení veličiny s tímto zpožděním do modelu ?** Je zřejmé, že připuštění zpoždění o značné délce (např. 10) mluví ve prospěch obecnosti modelu, na druhé straně může silně znesnadnit výpočet parametrů (nejsou-li tyto jinak specifikovány) u zpožděných veličin (s ohledem jednak na značnou pravděpodobnost výskytu přibližné kolinearity zpožděných veličin, jednak na malou hodnotu rozdílu T-k, jestliže je počet pozorování relativně malý a počet vysvětlujících proměnných naopak relativně velký) s obtížemi při testování jejich statistické významnosti.

b) **Jaká „váhová struktura" má být přiřazena jednotlivým zpožděným veličinám ?** Jen v málo případech je rozumné uvažovat stejnou sílu vlivu jednotlivých zahrnutých zpožděných vysvětlujících proměnných (v podstatě jde o rovnoměrné rozdělení vah) na závisle proměnnou. Daleko častější jsou situace, kdy k s rostoucí hloubkou zpoždění síla vlivu (značně) zpožděných veličin postupně klesá.

<u>**Poznámka :**</u> **I když v principu je model s nekonečnou délkou zpoždění nepříliš realistický, může být výpočet jeho parametrů (při konkrétně specifikované váhové struktuře) snadnější než komplikovaný model s konečně rozloženým zpožděním.**

Obecný tvar modelu s konečnou délkou zpoždění (k) je (při absenci konstantního členu regresní rovnice)

$$Y_t = \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + \dots + \beta_k X_{t-k} + \varepsilon_t \qquad \textbf{neboli}$$

$$Y_t = \sum_{j=0}^{k} \beta_j X_{t-j} + \varepsilon_t \qquad , \qquad t = 1,2,\dots,T$$

Každý z koeficientů β_j představuje váhu (též **koeficient reakce) j-té zpožděné veličiny na aktuální hodnotu proměnné X . O ε_t se předpokládají obvyklé vlastnosti náhodné složky lineárního regresního modelu.**

Některá konkrétní schémata modelů s konečným rozloženým zpožděním

1) Model s lineárně rozloženými zpožděními

Předpokládá se, že váhy w_j (označené zde β_j, neboť jde o regresní koeficienty) **u jednotlivých zpožděných vysvětlujících proměnných lineárně klesají od „nezpožděné" veličiny** X_t (nebo od X_{t-1} se zpožděním 1) **k veličině** X_{t-k} **s maximálním přípustným zpožděním k. To odpovídá předpokladu, že zpožděné proměnné vykazují slábnoucí vliv s konstantními úbytky.**

Poznámka : ***Ne nutně musí být součet přijatých vah roven 1*** *(normování na hodnotu 1 lze ostatně vždy zajistit),* **váhy však zpravidla budou všechny kladné. Znamená to, že váhovou strukturu lze popsat následovně :**

$$\beta_j = (k+1-j).\beta \qquad \textbf{pro} \quad j = 0,1,2,\ldots,k$$

$$\beta_j = 0 \qquad \textbf{pro} \quad j > k$$

Dosazením do základní rovnice dostaneme vztah

$$Y_t = \beta.\sum_{j=0}^{k}(k+1-j)X_{t-j} + \varepsilon_t = \beta.Z_t^{(k)} + \varepsilon_t \qquad , \qquad t = 1,2,\ldots.,T \quad ,$$

jehož jediný neznámý parametr β **může být snadno odhadnut MNČ jako**

$$\beta^* = \left(Z_k^{(k)}{}'.Z_k^{(k)}\right)^{-1} Z_k^{(k)}{}'Y_k = \frac{\sum Z_k^{(k)}.Y_k}{\sum Z_k^{(k)2}}$$

Z něj pak již snadno odvodíme odhady pro jednotlivá β_j **.**

2) Model s dvojlineárně rozloženými (střechovitými) zpožděními

Toto váhové schéma přikládá největší váhu hodnotě proměnné s určitým pevně zvoleným zpožděním (např. x_r **) , od něhož ve směru do přítomnosti i do minulosti váhy lineárně klesají. Předpokládáme-li , že** $\beta_0 = \beta_k = 0$ **, pak při sudém k můžeme strukturu zpoždění popsat těmito vztahy [1]*/ :**

$$\beta_j = j.\beta \qquad \textbf{pro} \quad j = 0,1,2,\ldots,k/2$$

$$\beta_j = (k-j).\beta \qquad \textbf{pro} \quad j = k/2, k/2+1,\ldots,k$$

Dosazením do základního modelu dostaneme regresní vztah

$$Y_t = \beta.\sum_{j=0}^{k/2} j.X_{t-j} + \beta.\sum_{j=k/2+1}^{k}(k-j).X_{t-j} + \varepsilon_t = \beta.Z_t^{(k)} + \varepsilon_t$$

jehož neznámý parametr β **lze opět odhadnut metodou OLS stejným způsobem jako v předchozím případě. Veličina** $Z_t^{(k)}$ **je zde definována jako**

$$Z_t^{(k)} = \sum_{j=0}^{k/2} j.X_{t-j} + \sum_{j=k/2+1}^{k}(k-j).X_{t-j}$$

Hodnoty původních parametrů β_j **získáme stejně jako u lineárního modelu.**

[1] **alternativně též** $\beta_j = (j+1) . \beta$ **pro j = 0, 1, 2, ... , k/2**
$\beta_j = (k-j+1) . \beta$ **pro j = k/2, k/2+1,.... , k**

3) Polynomický model Almonové [1965]

je příkladem modelu s konečnými rozdělenými zpožděními, u kterých je původní (početná) **množina parametrů převedena na** (počtem skromnější) **množinu jiných parametrů, přičemž příslušné transformační vztahy mezi parametry jsou předepsány maticovým schématem. Původní parametry se takto vyjádří jako lineární kombinace parametrů „nových" se zápisem**

$$\beta_{[k+1,1]} = M_{[k+1,r+1]}.\alpha_{[r+1,1]}$$

(původních *k+1* parametrů s označením β, nových *r+1* parametrů se značením α, transformační matice *M*) :

Konkrétně se v modelu Almonové předpokládá závislost polynomiálního tvaru, v němž je transformační vztah mezi parametry β a α popsán jako mnohočlen určitého (nevelkého) **stupně s formálním zápisem**

$$\beta_j = \alpha_0 + \alpha_1.j + \alpha_2.j^2 + \alpha_3.j^3 + \ldots. + \alpha_r.j^r, \qquad j = 0,1,2,\ldots.,k$$

Váhový vektor β zde můžeme vyjádřit kompaktněji maticovým zápisem :

$$\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \ldots. \\ \beta_k \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \ldots\ldots & 0 \\ 1 & 1 & 1 & \ldots\ldots. & 1 \\ 1 & 2 & 2^2 & \ldots\ldots.. & 2^r \\ \ldots. & \ldots\ldots & \ldots\ldots & \ldots\ldots. & \ldots\ldots \\ 1 & k & k^2 & \ldots\ldots.. & k^r \end{pmatrix} . \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ .\alpha_2 \\ \ldots\ldots \\ \alpha_r \end{pmatrix}$$

Původní regresní model tvaru $y = X\beta + \varepsilon$ lze při transformaci parametrů $\beta = M.\alpha$ přepsat do tvaru $y = X.M^{-1}.M\alpha + \varepsilon = Z.\alpha + \varepsilon$, takže původních $(k+1)$ váhových koeficientů je transformováno na menší počet $(r+1)$ parametrů . V transformovaném modelu spočteme MNČ odhad $\hat{\alpha}$, načež vektor odhadnutých vah zpoždění $\hat{\beta}$ získáme dosazením $\hat{\alpha}$ do vztahu $\hat{\beta} = M.\hat{\alpha}$.

<u>Příklad</u> se specifikací transformace jako polynomu 2.stupně :

$$\beta_j = \alpha_0 + \alpha_1.j + \alpha_2.j^2,$$

Dosazení do výchozího regresního váhového schématu vede k vyjádření :

$$Y_t = .\sum_{j=0}^{k}(\alpha_0 + \alpha_1.j + \alpha_2.j^2)..X_{t-j} + \varepsilon_t = \alpha_0.Z_{0t}{}^{(k)} + \alpha_1.Z_{1t}{}^{(k)} + \alpha_2.Z_{2t}{}^{(k)} + \varepsilon_t$$

, kde

$$Z_{0t}(k) = \sum_{j=0}^{k} X_{t-j} \qquad Z_{1t}(k) = \sum_{j=0}^{k} j.X_{t-j} \qquad Z_{2t}(k) = \sum_{j=0}^{k} j^2.X_{t-j}$$

Aplikujeme-li MNČ na tento model, spočteme odhady koeficientů $\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2$ **a po jejich dosazení do původního modelu získáme odhady výchozích váhových koeficientů** $\hat{\beta}_j$ **. V případě, že zvolíme maximální délku zpoždění např.** $k = 3$**, získáme po úpravě** (přeskupení proměnných dle příslušnosti ke koeficientům α_J) **pro vysvětlení závisle proměnné rovnici:**

$$Y_t = \alpha_0(X_t + X_{t-1} + X_{t-2} + X_{t-3}) + \alpha_1(X_{t-1} + 2X_{t-2} + 3X_{t-3})$$
$$+ \alpha_2(X_{t-1} + 4X_{t-2} + 9X_{t-3}) + \varepsilon_t$$

<u>**Poznámka :**</u> **počet** ***r*** **parametrů** $\mathbf{a_j}$ **je o jeden větší než stupeň polynomu, zatímco počet parametrů** β_j **je o jeden větší než volená hodnota maximálního zpoždění** k**.**

K tomu, abychom počet parametrů ještě případně dále omezili, můžeme využít „okrajových" podmínek $\beta_{-1} = \beta_{k+1} = 0$**, což vede k relacím :**

$$0 = \alpha_0 - \alpha_1 + \alpha_2 \qquad \text{při } j = -1$$
$$0 = \alpha_0 + \alpha_1(k+1) + \alpha_2(k+1)^2 \qquad \text{při } j = k+1$$

Z nich dostaneme (v závislosti na α_2 , které musí být záporné) :

$$\alpha_0 = -\alpha_2(k+1) \qquad , \qquad \alpha_1 = -\alpha_2.k\,.$$

Po substituci za α_0, α_1 lze takto konkretizovaný model psát jako

$$Y_t = \alpha_2.Z_t^{(k)} + \varepsilon_t \quad \text{, \textbf{přičemž}}$$

$$Z_t^{(k)} = \sum_{j=0}^{k} (j^2 - k - k.j - 1)X_{t-j}$$

Připomeňme, že platí

$$Y_t = \alpha_2.Z_t^{(k)} + \varepsilon_t \quad \text{, \textbf{přičemž}}$$

$$Z_t^{(k)} = \sum_{j=0}^{K} \alpha_2(-(k+1) - kj + j^2)X_{t-j}$$

<u>**Pro**</u> **náš** <u>**konkrétní případ k=3 dostaneme**</u>**:** $\alpha_1 = -3\alpha_2$ **,** $\alpha_0 = -4\alpha_2$

Výraz $\sum_{j=0}^{k}(j^2 - k(j+1) - 1)$ **je pro** $k = 3$ **roven** $\sum_{j=0}^{3}(j^2 - 3j - 4)$

a koeficienty u jednotlivých zpožděných proměnných jsou následující

$Y_t = -4\alpha_2.X_t - 6\alpha_2 X_{t-1} - 6\alpha_2.X_{t-2} - 4\alpha_2.X_{t-3}$ (při záporném α_2) , resp.

$$Y_t = -2\alpha_2.\left(2X_t + 3X_{t-1} + 3X_{t-2} + 2X_{t-3}\right).$$

Odhady parametrů získané tímto postupem jsou vydatnější než odhady získané přímo metodou OLS. Nevýhodou postupu je, že zpravidla nemáme vodítka po určení maximální délky zpoždění. Doporučuje se použít pro zvolený polynom nízkého řádu (max.3-4) různé délky max. zpoždění a vybrat tu, u které je hodnota korigovaného R^2 největší. Určení stupně polynomu je nicméně **obvykle subjektivní záležitostí.**

Při určení maximální délky zpoždění bude hrát roli mj. též povaha vzorku časových řad – u čtvrtletních či měsíčních bude mít maximální zpoždění větší hodnotu než v případě odhadu modelu z ročních časových řad.

4) <u>Koyckův model - Koyckova transformace</u>

je (naopak) **příkladem modelu s rozloženým zpožděním o nekonečné délce. Má-li být zachována možnost statisticky odhadnout parametry takovýchto modelů, musí být dáno nějaké pravidlo o souvislostech mezi nimi. V případě modelu navrženého Koyckem klesají váhy u jednotlivých vysvětlujících zpožděných proměnných podle schématu popsaného geometrickou posloupností.**
Zapíšeme-li základní rovnici modelu s nekonečně rozloženým zpožděním ve tvaru

$$Y_t = \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + \ldots. + \beta_k X_{t-k} + \ldots\ldots\ldots + \varepsilon_t$$

neboli ve zkráceném zápisu $Y_t = \sum_{j=0}^{\infty} \beta_j X_{t-j} + \varepsilon_t$

je ihned patrné, že takto obecně vyjádřený model nelze prakticky použít (nelze odhadnout nekonečný počet parametrů)**. Dle Koyckem navržené konkretizace přijímají parametry tuto apriorní váhovou strukturu :**

$$\beta = b.w_j \qquad j = 0,1,2,\ldots.,$$ **kde**

váhy/koeficienty w_j jsou prvky geometrické posloupnosti

$$w_j = (1-q).q^j, \qquad 0 < q < 1 \qquad ,$$

která je pro danou hodnotu kvocientu q **klesající.**

Tímto způsobem lze převést původně nekonečný počet parametrů pouze na dva parametry b a **q , přičemž v konečné podobě model nabude tvar**

$$Y_t - q.Y_{t-1} = (1-q)\, X_t + (\varepsilon_t - q.\varepsilon_{t-1}) ,$$ **což lze upravit na tvar**

$$\mathbf{Y_t = q.Y_{t-1} + b.(1\text{-}q)\, X_t + v_t} \qquad ,$$

který je nazýván autoregresním tvarem modelu (nekonečného) rozloženého zpoždění. Všimněme si zde zejména dvou věcí :

a) do modelu se na pravou stranu dostala (jediná) zpožděná <u>závisle</u> proměnná (se zpožděním o 1 krok)

b) náhodné složky modelu $v_t = \varepsilon_t - q.\ \varepsilon_{t-1}$ již (bohužel) **nebudou vzájemně nekorelované, a to ani tehdy ne, jestliže byla předpokládána nekorelovanost původních náhodných složek ε_t. Příčinou toho je skutečnost, že „nová" vysvětlující proměnná Y_{t-1} není nekorelovaná s náhodnými složkami v_t. Platí totiž :**

$$E[Y_{t-1}.(\varepsilon_t - q.\ \varepsilon_{t-1})] = E[\varepsilon_{t-1}.(\varepsilon_t - q.\varepsilon_{t-1})] = E[\varepsilon_{t-1}.\ \varepsilon_t] - E[\varepsilon_{t-1}(q.\ \varepsilon_{t-1})] = -q.\sigma^2$$

kde σ^2 je rozptyl náhodných složek ε_t . Oproti klasickému lineárnímu regresnímu modelu tedy zde zřejmě nejsou splněny dva předpoklady :

- vysvětlující proměnná Y_{t-1} není nekorelovaná s náhodnou složkou v_t
- vysvětlující proměnná Y_{t-1} není nestochastická (její součástí je náhodná složka u_{t-1}), což je hned vidět, zapíšeme-li model se zpožděním o 1 krok.

Z uvedených důvodů nemohou mít odhady parametrů (provedené obyčejnou metodou nejmenších čtverců) uspokojivé vlastnosti, nemusí být dokonce ani konzistentní. Literatura uvádí pro tuto a podobné situace některé speciální odhadové postupy (vedoucí ke konzistentním, případně i vydatným odhadům parametrů). Předpoklady o chování náhodných složek podmiňující nasazení těchto postupů jsou však obvykle málo realistické .

S ohledem na vlastnosti geometrického rozdělení přijatého v Koyckově modelu činí průměrná délka zpoždění hodnotu $q/(1-q)$ a rozptyl $q/(1-q)^2$.

Při q = 1/3 bude EX = 1/3 : 2/3 = 1/2
Při q = 1/2 bude EX = 1/2 : 1/2 = 1
Při q = 2/3 bude EX = 2/3 : 1/3 = 2

Interpretačně to znamená, že agregovaný účinek všech v modelu uvažovaných zpožděných vysvětlujících veličin (jichž je nekonečně mnoho) se projeví zhruba stejně jako jediná zpožděná vysvětlující proměnná, která bude mít zpoždění 0,5 roku resp. 1 rok, resp. 2 roky.

Podrobněji bude Koyckův model uveden v Modely RZ2