KONEČNĚ ROZDĚLENÁ ZPOŽDĚNÍ. POLYNOMICKY ROZDĚLENÉ ZPOŽDĚNÍ.

Tento text je zaměřen na modely konečně zpoždění, podrobněji je pak rozebráno polynomicky rozdělené zpoždění. Občas bývá rozumné zahrnout do modelu nejen současné, ale i minulé hodnoty určitých proměnných. Například pokud předpokládáme, že dovoz do země závisí nejen na HDP v současném období, ale i na HDP v minulém období, nebo na dovozu v minulém období, nebo na obojím, je vhodné to v modelu nějak zohlednit. Z modelu **statického** se tak stane model **dynamický**.

V případě, že do modelu zahrneme zpožděné hodnoty vysvětlující proměnné, nazývá se to **model rozdělených zpoždění** (distributed lag). Pokud zahrneme zpožděné hodnoty vysvětlované proměnné, jde o **autoregresní model** (autoregressive mode). A konečně, pokud zahrneme obojí, říká se tomu autoregressive-distributed-lag, ADL model.

Musíme však nějakým způsobem určit **délku zpoždění** a jeho **strukturu**. Na to neexistuje striktní postup, můžeme vyjít z ekonomické hypotézy či si pomoci informačními kritérii (Akaikeho AIC, Hannan-Quinn, Schwarz), a hledáme model, kde budou tato kritéria co nejmenší. Pro určení struktury zpoždění potřebujeme mít k dispozici dodatečnou informaci, jak by to nejspíš mohlo vypadat, tedy jestli by třeba koeficienty mohly trvale klesat apod. Často můžeme předpokládat, že váhy jsou generovány nějakou známou pravděpodobnostní funkcí.

Máme-li nějaké informace o délce zpoždění, můžeme použít modely **konečně rozděleného zpoždění**, kdy do modelu zahrneme třeba jen dvě nebo tři zpožděné proměnné. Zpožděné vlivy vysvětlujících proměnných tak trvají pouze určitý čas a pak jejich vliv vyprchá.

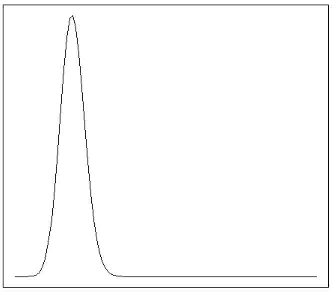
Co se týče struktury koeficientů, čili určení vah parametrů, můžeme předpokládat:

- že trvale rovnoměrně klesají (druhé pozorování má menší vliv než první, třetí než druhé atd.) → **aritmeticky rozdělené zpoždění**
- že váhy nejprve lineárně rostou a pak lineárně klesají → „**obrácené V**“
- že jejich forma není lineární (klesají, rostou, klesají…) → **aproximace polynomem nízkého stupně**

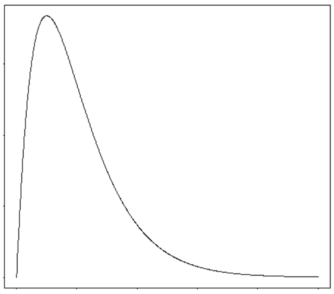
Jestliže nemáme informace o délce zpoždění, vyjdeme většinou z **modelu nekonečně rozdělených zpoždění**. Opět je potřeba určit, jak mohou vypadat váhy parametrů (reakční koeficienty) zpožděných vysvětlujících proměnných.

- nejčastěji předpokládáme, že klesají geometrickou řadou → geometrické zpoždění
- také si ale můžeme myslet, že maximální efekt má jiná vysvětlující proměnná než ta z posledního období → Pascalovo rozdělení
- další možnosti: racionálně rozdělené zpoždění, gama rozdělené zpoždění, exponenciálně rozdělené zpoždění…

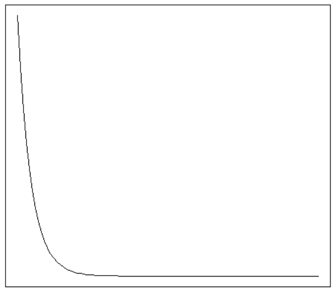
Pascalovo rozdělení



Gamma rozdělení



Geometrické rozdělení



Modely nekonečně rozděleného zpoždění jsou podrobněji rozebrány v otázce „Modely rozdělených zpoždění. Friedmanova spotřební funkce a permanentní důchod.“

KONEČNĚ ROZDĚLENÁ ZPOŽDĚNÍ

Obecný tvar modelu konečně rozdělených zpoždění je následující:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} \ldots \beta_n X_{t-n} + u_t$$

To znamená, že současná hodnota vysvětlované proměnné je závislá na současných a minulých hodnotách vysvětlujících proměnných. Při splnění GM předpokladů můžeme takovýto model odhadnout metodou nejmenších čtverců. Parametry β mají následující interpretaci.

Představme si, že X bylo konstantní a pak se v období t zvýšilo o jednotku, ale v dalších obdobích se vrátilo zpět na svou původní úroveň.

- ➔ β_0 nám říká, jak se v průměru změní Y_t. Říká se mu **běžný multiplikátor**.
- ➔ β_1 nám říká, jak se v průměru změní Y_{t+1}. Říká se mu **dynamický multiplikátor** (zpožděný o jedno období).
- ➔ β_i nám říká, jak se v průměru změní Y_{t+i}. Říká se mu dynamický multiplikátor (zpožděný o i období).

Představme si, že X bylo konstantní a pak se v období t zvýšilo o jednotku a už zůstalo vyšší.

- ➔ β_0 nám říká, jak se v průměru změní Y_t. Říká se mu běžný multiplikátor.
- ➔ $\beta_0 + \beta_1$ nám říká, jak se v průměru změní Y_{t+1}. Říká se mu **střednědobý kumulativní multiplikátor**.
- ➔ $\beta_0 + \beta_1 + \ldots \beta_i$ nám říká, jak se v průměru změní Y_{t+i}. Říká se mu střednědobý kumulativní multiplikátor.
- ➔ $\sum_{i=0}^{n} \beta_i$ nám říká, jaká bude celková průměrná změna Y_{t+q}. Říká se mu **celkový multiplikátor**.

Co když se ale netrefíme s délkou zpoždění? Jestliže je skutečné zpoždění n*, pak

a) při specifikaci $n < n^*$ nebudou kvůli vypuštění relevantních vysvětlujících proměnných odhady parametrů nestranné

b) při $n > n^*$ nebudou kvůli přítomnosti irelevantních proměnných odhady parametrů vydatné

Příklad: podívej se na následující odhad modelu, pomocí něhož šikovný ekonometr jménem Bob odhadl závislost své měsíční útraty za kulturu (Y) na výši svého měsíčního platu (X) s požitím modelu konečně rozdělených zpoždění.

$$Y_t = 50 + 0{,}08X_t + 0{,}04X_{t-1} + 0{,}02X_{t-2} + 0{,}01X_{t-3}$$

čas	běžný multiplikátor	dynamické multiplikátory	střednědobé kumulativní multiplikátory	celkový multiplikátor
t = 0	0,08			
t = 1		0,04	0,12	
t = 2		0,02	0,14	
t = 3		0,01		0,15

Při své běžné výši platu, která činí 10 000 Kč (Bob pracuje na part-time, ne že by si ekonometrové vydělávali tak málo, zároveň ještě studuje), Bob průměrně utratí za kulturu 1 550 Kč měsíčně.

*Když Bob dostane v prosinci bonus ke své výplatě ve výši 1 000 Kč, tak v prosinci utratí za kulturu v průměru o 80 Kč více = **běžný multiplikátor**.*

*V lednu už Bob žádný bonus nedostane, protože práci fláká, neboť se pilně připravuje na zkoušky. Přesto i v lednu utratí díky prosincovému měsíčnímu bonusu v průměru o 40 Kč více za kulturu = **dynamický multiplikátor** zpožděný o jedno období.*

Kdyby ale Bob dostal bonus ve výši 1 000 Kč jak v prosinci, tak v lednu, utratil by v lednu za kulturu v průměru o 120 Kč více než obvykle = střednědobý kumulativní multiplikátor.

*Bobův šéf se rozhodl, že mu trvale zvýší plat o 1 000 Kč. Díky tomu bude Bob v průměru utrácet o 150 Kč více za kulturu než dříve = **celkový multiplikátor**.*

Přesné odhady parametrů může být těžké získat, protože mezi zpožděnými hodnotami vysvětlující proměnné bývá silná kolinearita, v modelu se často vyskytuje autokorelace, a směrodatné chyby odhadnutých parametrů bývají tedy vysoké (to znamená, že odhady mají velkou variabilitu a při jiném výběru můžou vyjít úplně jinak).

Může ale pomoci to, když do modelu zahrneme **apriorní informaci** o tvaru zpoždění. To znamená, že přijmeme určitý předpoklad o tom, jak nejspíš působí vysvětlující proměnná na vysvětlovanou (zda její vliv s rostoucím časem stále klesá, nebo nejprve roste a pak klesá apod.). Na základě tohoto předpokladu zavedeme určitou restrikci parametrů a odhadneme tzv. omezený model, z něhož pak můžeme parametry původního modelu zpětně dopočítat.

ARITMERICKY ROZDĚLENÉ ZPOŽDĚNÍ:

Jestliže předpokládáme, že váhové koeficienty klesají aritmetickou řadou, můžeme je definovat jako

$\beta_i = (n + 1 - i)\beta$ pro $i = 0,1 \ldots n,$
$\beta_i = 0$ pro $i > n.$

Dosadíme-li to do vztahu $Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} \ldots \beta_q X_{t-q} + u_t$, dostaneme:
$Y_t = \beta \sum_{i=0}^{n}(n+1-i)X_{t-i} + u_t \rightarrow Y_t = \beta Z_t + u_t$

Například pro n = 3 (vliv mají 3 zpožděné hodnoty vysvětlující proměnné) platí:
$\beta_0 = (3 + 1 - 0)\beta = 4\beta;$
$\beta_1 = (3 + 1 - 1)\beta = 3\beta;$
$\beta_2 = 2\beta;$
$\beta_3 = 1\beta;$
$\beta_4 = 0;$
$\beta_5 = 0; \ldots$

$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + u_t$
$Y_t = \alpha + 4\beta X_t + 3\beta X_{t-1} + 2\beta X_{t-2} + 1\beta X_{t-3} + u_t$
$Y_t = \alpha + \beta(4X_t + 3X_{t-1} + 2X_{t-2} + 1X_{t-3}) + u_t$
$Y_t = \alpha + \beta Z_t + u_t,$ kde $Z_t = 4X_t + 3X_{t-1} + 2X_{t-2} + 1X_{t-3}$

Model $Y_t = \alpha + \beta Z_t + u_t$ je snadné odhadnout klasickou MNČ. Tak získáme β a dopočítáme β_i.

OBRÁCENÉ V

Jestliže předpokládáme, že váhové koeficienty klesají aritmetickou řadou, můžeme je definovat jako
$\beta_i = (i + 1)\beta$ pro $0 \le i \le n/2$
$\beta_i = (n + 1 - i)\beta$ pro $i \ge 1 + n/2$

Dosadíme-li to do vztahu $Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} \ldots \beta_q X_{t-q} + u_t$,, dostaneme:
$Y_t = \beta[\sum_{i=0}^{n/2}(i+1)X_{t-i} + \sum_{i=1+n/2}^{n}(n+1-i)X_{t-i}] + u_t \rightarrow Y_t = \beta Z_t + u_t$

Například pro n = 4 (vliv mají 4 zpožděné hodnoty vysvětlující proměnné) platí:
$\beta_0 = (0 + 1)\beta = 1\beta;$
$\beta_1 = (1 + 1)\beta = 2\beta;$
$\beta_2 = (2 + 1)\beta = 3\beta;$
$\beta_3 = (4 + 1 - 3) = 2\beta;$
$\beta_4 = (4 + 1 - 4) = 1\beta;$
$\beta_5 = 0; \ldots$

$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + \beta_4 X_{t-4} + u_t$
$Y_t = \alpha + 1\beta X_t + 2\beta X_{t-1} + 3\beta X_{t-2} + 2\beta X_{t-3} + 1\beta_4 X_{t-4} + u_t$
$Y_t = \alpha + \beta(1X_t + 2X_{t-1} + 3X_{t-2} + 2X_{t-3} + 1X_{t-4}) + u_t$
$Y_t = \alpha + \beta Z_t + u_t,$ kde $Z_t = 1X_t + 2X_{t-1} + 3X_{t-2} + 2X_{t-3} + 1X_{t-4}$

Opět, model $Y_t = \alpha + \beta Z_t + u_t$ je snadné odhadnout klasickou MNČ. Tak získáme β a dopočítáme β_i.

POLYNOMICKY ROZDĚLENÉ ZPOŽDĚNÍ

Jestliže by nám pomohlo specifikovat pružnější formu váhových koeficientů, můžeme využít model Almonové. Tento model, nazývaný model polynomicky rozdělných zpoždění, předpokládá, že průběh vah v čase se dá aproximovat polynomem nízkého řádu, obecně:

$$\beta_i = a_0 + a_1 i + a_2 i^2 + \ldots + a_r i^r$$

Jak vybrat polynom? Pro tento polynom by mělo platit, že je menší než délka zpoždění, ale větší než počet bodů zvratů křivky. Když váhy rostou a pak klesají (jako je tomu u obráceného V), tak je dobrý polynom 2. řádu. Jestliže má trajektorie β cyklický charakter, tak je dobrý polynom 3. řádu.

Příklad bude ukázán pro polynom 2. řádu. Pak tedy apriorním omezením vah bude:

$\beta_i = a_0 + a_1 i + a_2 i^2$ pro $i = 0,1\ldots n$ (kde n je délka zpoždění)

Tedy: $Y_t = \alpha + a_0 X_t + (a_0 + a_1 + a_2)X_{t-1} + \ldots (a_0 + na_1 + n^2 a_2)X_{t-n} + u_t$

$Y_t = \alpha + a_0 \sum_{i=0}^{n} X_{t-i} + a_1 \sum_{i=0}^{n} iX_{t-i} + a_2 \sum_{i=0}^{n} i^2 X_{t-i} + u_t$

Definují se tři nové proměnné: $Z_{0t} = \sum_{i=0}^{n} X_{t-i}$; $Z_{1t} = \sum_{i=0}^{n} iX_{t-i}$; $Z_{2t} = \sum_{i=0}^{n} i^2 X_{t-i}$

A přepíšeme model jako $Y_t = \alpha + a_0 Z_{0t} + a_1 Z_{1t} + a_2 Z_{2t} + u_t.$

To je **omezený** model konečně rozděleného zpoždění, protože *n* + 1 neznámých váhových koeficientů je redukováno jen na *r* + 1 koeficientů (zde pro *r* = 2 máme 3 koeficienty). Zpožděné hodnoty *X* jsou nahrazeny jejich lineárními kombinacemi (proměnnými *Z*).

Příklad pro n = 4:

$\beta_i = a_0 + a_1 i + a_2 i^2$
$b_0 = a_0 + 0a_1 + 0a_2 = a_0$
$b_1 = a_0 + 1a_1 + 1a_2$
$b_2 = a_0 + 2a_1 + 4a_2$
$b_3 = a_0 + 3a_1 + 9a_2$
$b_4 = a_0 + 4a_1 + 16a_2$

$Z_{0t} = X_t + X_{t-1} + X_{t-2} + X_{t-3} + X_{t-4}$
$Z_{1t} = 1X_{t-1} + 2X_{t-2} + 3X_{t-3} + 4X_{t-4}$
$Z_{2t} = 1X_{t-1} + 4X_{t-2} + 9X_{t-3} + 16X_{t-4}$

Proměnné *Z* si tedy dopočítáme z *X*, odhadneme a_0, a_1, a_2 a dopočítáme *b*.
Výhodou je snížení multikolinearity.
Problémem v praxi bývá neznalost délky zpoždění *n*. Chce to vyzkoušet a porovnat modely s různou délkou zpoždění například pomocí informačních kritérií. Zároveň ale neznáme ani ten „nejlepší“ polynom. Ideální je udělat to dvoustupňově, tedy určit si nejprve délku zpoždění, a pak pro danou délku určit polynom.

PŘÍKLAD

(Zdroj: 9. cvičení ECON2300 University of Queensland)

Následující výstup zachycuje odhad modelu závislosti počtu akrů pěstované cukrové třtiny na farmě pana Donalda (A) na ceně cukrové třtiny (P) v Bangladéši. Konkrétně jde o následující model (model 1):

$$\ln(A_t) = \alpha + \beta_0 \ln(P_t) + \beta_1 \ln(P_{t-1}) + \beta_2 \ln(P_{t-2}) + \beta_3 \ln(P_{t-3}) + \beta_4 \ln(P_{t-4}) + u_t$$

Pracujeme tedy s proměnným ve tvaru logaritmů a jde o model konečně rozdělených zpoždění, kdy má vliv současná a čtyři předchozí hodnoty vysvětlující proměnné. Odhad modelu v E-views vypadá následovně:

Dependent Variable: LOG(A)
Method: Least Squares
Date: 05/10/09 Time: 13:47
Sample (adjusted): 5 34
Included observations: 30 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.824119	0.100580	38.02070	0.0000
LOG(P)	0.774623	0.312860	2.475943	0.0207
LOG(P(-1))	-0.217455	0.318546	-0.682648	0.5014
LOG(P(-2))	-0.002556	0.322066	-0.007937	0.9937
LOG(P(-3))	0.586819	0.315268	1.861338	0.0750
LOG(P(-4))	-0.014325	0.298523	-0.047988	0.9621

R-squared	0.298054	Mean dependent var	3.979527
Adjusted R-squared	0.151816	S.D. dependent var	0.331706
S.E. of regression	0.305491	Akaike info criterion	0.643063
Sum squared resid	2.239792	Schwarz criterion	0.923302
Log likelihood	-3.645942	Hannan-Quinn criter.	0.732714
F-statistic	2.038136	Durbin-Watson stat	1.083112
Prob(F-statistic)	0.109198		

Zkuste si na následující otázky nejprve odpovědět sami:

1) Pozor na to, jak se interpretují zlogaritmované proměnné. Konkrétně zde se parametr β_0 interpretuje tak, že:
   a) když se současná cena zvedne o jednotku, zvedne se počet akrů také o 0,77 jednotek
   b) když se současná cena zvedne o jednotku, zvedne se počet akrů o 0,77 %
   c) když se současná cena zvedne o 1 %, zvedne se počet akrů o 0,77 jednotek
   d) když se současná cena zvedne o 1 %, zvedne se počet akrů o 0,77 %
2) Jaký podíl variability vysvětlované proměnné se podařilo modelem vysvětlit? ____________
3) Kolik parametrů je dle modelu, nepočítáme-li úrovňovou konstantu, statisticky významných na 5% hladině významnosti? __________
4) Co nejspíš způsobilo statistickou nevýznamnost některých parametrů? ____________

Odpovědi:

1) *Správně je d), protože vysvětlující i vysvětlovaná proměnná jsou ve tvaru logaritmů.*
2) *Podařilo se nám vysvětlit 29,8 % variability vysvětlované proměnné. Zjistíme to z koeficientu determinace.*
3) *Pouze jeden (β_0), jehož p-hodnota je menší než 0,05.*

4) *Příčinou je pravděpodobně multikolinearita. Je těžké modelem zachytit oddělený vliv jednotlivých proměnných. Pomoci by mohlo to, kdybychom využili apriorní informaci o vahách parametrů. Pak bychom mohli sestavit omezený model s menším počtem proměnných, což by mohlo vyřešit problém multikolinearity.*

Model není nic moc, usoudil farmář Donald. Odhadl tedy nový model s apriorním omezením vah. Rozhodl se, že použije aproximaci vhodným polynomem. Délku zpoždění nechal opět stejnou, čili n = 4, a a β_i aproximoval vztahem:

$$\beta_i = \alpha_0 + i\alpha_1$$

Zde je model (model 2) a jeho odhad:

$$\ln(A_t) = \alpha + \alpha_0 z_{t0} + \alpha_1 z_{t1} + u_t, \text{ kde}$$
$$z_{t0} = \ln(P_t) + \ln(P_{t-1}) + \ln(P_{t-2}) + \ln(P_{t-3}) + \ln(P_{t-4})$$
$$z_{t1} = \ln(P_{t-1}) + 2\ln(P_{t-2}) + 3\ln(P_{t-3}) + 4\ln(P_{t-4})$$

Dependent Variable: LOG(A)
Method: Least Squares
Date: 05/10/09 Time: 14:12
Sample (adjusted): 5 34
Included observations: 30 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.822599	0.105617	36.19311	0.0000
Z0	0.424668	0.259444	1.636838	0.1133
Z1	-0.099630	0.108812	-0.915617	0.3680

R-squared	0.122572	Mean dependent var	3.979527
Adjusted R-squared	0.057578	S.D. dependent var	0.331706
S.E. of regression	0.322015	Akaike info criterion	0.666201
Sum squared resid	2.799726	Schwarz criterion	0.806321
Log likelihood	-6.993020	Hannan-Quinn criter.	0.711027
F-statistic	1.885883	Durbin-Watson stat	1.426373
Prob(F-statistic)	0.171141		

Čemu se dle tohoto modelu budou rovnat parametry β_i?

$\beta_0 = \alpha_0 + 0\alpha_1 = 0{,}42$
$\beta_1 = \alpha_0 + 1\alpha_1 = 0{,}32$
$\beta_2 = \alpha_0 + 2\alpha_1 = 0{,}22$
$\beta_3 = \alpha_0 + 3\alpha_1 = 0{,}12$
$\beta_4 = \alpha_0 + 4\alpha_1 = 0{,}02$

Odhadnuté modely tedy vypadají následovně:

$\ln(\hat{A}_t) = \alpha + \beta_0 \ln(P_t) + \beta_1 \ln(P_{t-1}) + \beta_2 \ln(P_{t-2}) + \beta_3 \ln(P_{t-3}) + \beta_4 \ln(P_{t-4})$

Model 1: $\ln(\hat{A}_t) = 3{,}82 + 0{,}77 \ln(P_t) - 0{,}21 \ln(P_{t-1}) - 0{,}003 \ln(P_{t-2}) + 0{,}59 \ln(P_{t-3}) - 0{,}01 \ln(P_{t-4})$

Model 2: $\ln(\hat{A}_t) = 3{,}82 + 0{,}42 \ln(P_t) + 0{,}32 \ln(P_{t-1}) + 0{,}22 \ln(P_{t-2}) + 0{,}12 \ln(P_{t-3}) + 0{,}02 \ln(P_{t-4})$

Ani druhý model není nikterak skvělý. Nicméně vidíme, že odhadnuté parametry mají hodnoty smysluplnější hodnoty. Druhý model získaný s využitím apriorní informace o vahách parametrů bychom preferovali před prvním.

ZDROJE

Hušek, R: Ekonometrická analýza. Nakladatelství Oeconomica, Praha 2007.

Berry, K., Wei, S.: Distributed Lag Models. Unversity of Wyoming, 2013.
http://www.uwyo.edu/aadland/classes/econ5360/notes2.pdf

University of Queensland, ECON2300 (lecture 9 and tutorial 9), Australia 2012.