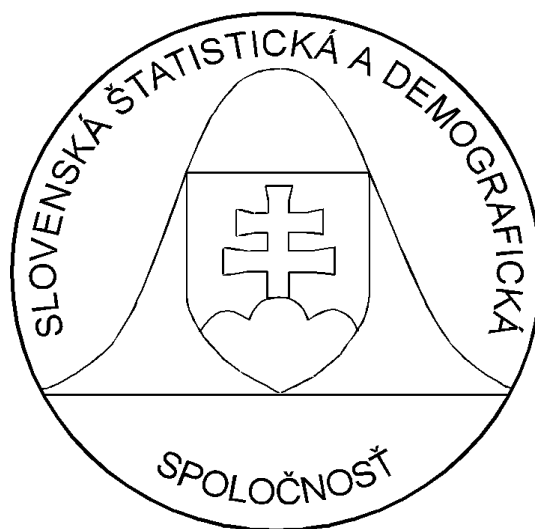


1/2007

FORUM STATISTICUM SLOVACUM



ISSN 1336-7420



9771336742001

7 1



**Slovenská štatistická a demografická
spoločnosť Miletičova 3, 824 67
Bratislava
www.ssds.sk**



Naše najbližšie akcie:

(pozri tiež www.ssds.sk, blok Poriadané akcie)

11. SLOVENSKÁ DEMOGRAFICKÁ KONFERENCIA,

tematické zameranie: *Migrácia*

17. – 19. 9. 2007, hotel Čingov, Slovenský raj

FernStat 2007

IV. medzinárodná konferencia aplikovanej štatistiky

(Financie, Ekonomika, Riadenie, Názory)

tematické zameranie: *Aplikovaná, demografická, matematická štatistika,
štatistické riadenie kvality.*

4. – 5. 10. 2007, hotel Lesák, Tajov pri Banskej Bystrici

16. Medzinárodný seminár VÝPOČTOVÁ ŠTATISTIKA,

6. – 7. 12. 2007, Bratislava

Prehliadka prác mladých štatistikov a demografov

6. 12. 2007, Bratislava

SLÁVNOSTNÁ KONFERENCIA 40 ROKOV SŠDS,

27. 3. 2008, Bratislava

KONFERENCIA POHĽADY NA EKONOMIKU SLOVENSKA 2008,

tematické zameranie: *Vývoj HDP a dlhodobá nezamestnanosť*

15. 4. 2008, Bratislava

EKOMSTAT 2007, 22. škola štatistiky

Tematické zameranie: *Štatistické metódy v praxi*

jún 2008, Trenčianske Teplice

14. SLOVENSKÁ ŠTATISTICKÁ KONFERENCIA,

tematické zameranie: *Regionálna štatistika*

rok 2008, Žilinský kraj

ÚVOD

Vážené kolegyně, vážení kolegovia,
prvé číslo tretieho ročníka časopisu, ktorý vydáva Slovenská štatistická a demografická spoločnosť (SŠDS) je zostavené z príspevkov, ktoré autori pripravili pre 21. školu štatistiky EKOMSTAT 2007 s tematickým zameraním „Štatistické metódy v praxi“. Táto akcia sa uskutočnila v dňoch 3. – 8. júna 2007 v Domove speváckeho zboru slovenských učiteľov v Trenčianskych Tepliciach.

Akciu, z poverenia Výboru SŠDS, zorganizoval Organizačný a programový výbor: Doc. Ing. Jozef Chajdiak, CSc. – predseda, RNDr. Ján Luha, CSc. – tajomník, RNDr. Samuel Koróny, Doc. RNDr. Viliam Páleník, PhD, Ing. Vladimír Kvetan.

Na príprave a zostavení tohoto čísla participovali: Doc. Ing. Jozef Chajdiak, CSc., RNDr. Ján Luha, CSc.

Recenziu príspevkov zabezpečili: Doc. Ing. Jozef Chajdiak, CSc., RNDr. Ján Luha, CSc., RNDr. Samuel Koróny.

Výbor SŠDS

Kvótový výber

RNDr. Ján Luha, CSc.

ABSTRACT

The article deals with quota sample methods. It describes procedure for construction of quota sample for application of face-to-face interview through the means of interviewers' network. It illustrates computation of quotas and specification of figures for the number of interviews per network. It also gives reference to some arguments to be found in literature that support exploitation of formulas concerning random sample for the quota sample.

1. ÚVOD

Kvótový výber sa stále často využíva v empirických prieskumoch kedy potrebujeme získať reprezentatívnu výberovú vzorku. V príspevku sa venujeme definícii výberového súboru všeobecne, dôraz kladieme na definíciu a vlastnosti kvótového výberu, na konštrukciu kvótového výberu vrátane konštrukcie rozpisu kvót na anketársku sieť. Často sa stretávame s argumentami za i proti realizácii zberu dát pomocou kvótového výberu, preto sa v príspevku do určitej miery zaoberáme aj argumentami, ktoré prirovnávajú kvótový výber ku modelu oblastného (stratifikovaného) náhodného výberu a tým argumentujú aj možnosť aplikovať na výpočet výberových odhadov známe vzorce z teórie oblastných náhodných výberov.

Samotný proces výberového zisťovania možno rozdeliť, v závislosti od cieľov skúmania, do niekoľkých fáz, napr.:

1. Definícia záujmového základného súboru. Napr. populácia SR vo veku 18 a viac rokov.
2. Špecifikácia opory výberu (pre náhodný výber), prípadne základných parametrov kontrolovaných kvótových znakov (pre kvótový výber).
3. Určenie metódy výberu a rozsahu výberu. Obvykle náhodný náhodný alebo kvótový výber.
4. Realizácia zberu dát.
5. Zhodnotenie realizácie výberového procesu.

Poznámka: V príspevku sa budeme zaoberať metódou kvótového výberu. Kvôli úplnosti stručne uvedieme aj základné všeobecné informácie o výberových metódach.

2. VÝBEROVÉ ŠTATISTICKÉ METÓDY

2.1 Všeobecná definícia výberu, základné pojmy

Výberom je vo všeobecnosti ľubovoľná podmnožina základného súboru. Nech Z označuje množinu prvkov záujmového základného súboru, potom jeho ľubovoľná podmnožina V je výberom zo Z .

Všeobecne teda definujeme výber: $V \subset Z$.

Špecifikáciou vlastností na metódu výberu prvkov dostávame konkrétne typy výberových metód.

Hlavným cieľom výberových zisťovaní je získať **reprezentatívny** výber zo záujmového základného súboru. Princíp reprezentatívnosti možno jednoducho parafrázovať výrokom Georga Gallupa: Ak kuchárka dobre zamieša obsah hrnca, tak jej stačí nabrat' a ochutnať iba za lyžicu polievky, aby zistila, ako to vyzerá s obsahom celého hrnca.

Výberové štatistické metódy dávajú teoretický základ ku skúmaniu rozsiahlych súborov. Na zabezpečenie kvalitnej kvantitatívnej analýzy, ktorá nám dá obraz o stave skúmaného základného súboru (napríklad súbor dospelej populácie SR, súbor priemyselných podnikov SR, súbor klientov banky, súbor poistencov zdravotnej poisťovne, súbor poľnohospodárskych podnikov) musíme získať kvalitné údaje.

Zber dát je najdelikátnejšou a najnákladnejšou operáciou a ani najpremyslenejšie metódy matematickej štatistiky naznamenajú nič, ak sa uplatňujú na informáciách zaťažených veľkými chybami.

Teória výberových zisťovaní má snahu poskytnúť metodologickú oporu pre nájdenie najprimeranejších prostriedkov zberu informácií.

Pri štatistickom zisťovaní býva najčastejšou voľba medzi dvoma stratégiami:

- skúmať všetky jednotky súboru, t.j. vyčerpávajúce (úplné) zisťovanie
- skúmať (podľa určitého plánu) vzorku jednotiek, t.j. výberové zisťovanie, na základe ktorého vyvodzujeme úsudky na celý súbor.

Ako príklady vyčerpávajúcich zisťovaní poznáme sčítania obyvateľov, domov a bytov, poľnohospodárskych zvierat a i.

V určitých prípadoch nemáme inú možnosť ako je výberové zisťovanie, napr. keď:

- skúmanie vedie ku uničeniu, alebo spotrebovaniu výrobku (pri kontrole jeho kvality)
- skúmaný základný súbor je prakticky neprístupný (napr. skúmanie maloobchodných cien)
- je náš rozpočet obmedzený.

Ak si máme možnosť vybrať, slúžia ako kritériá **náklady a kvalita** (presnosť) získaných dát.

Náklady na výberové zisťovanie sú spravidla nižšie než na úplné zisťovanie.

Kvalita môže byť u výberového skúmania, ktoré je dobre pripravené a dobre realizované dokonca väčšia a teda s presnejšími výsledkami. To súvisí s dvoma skupinami chýb:

- chyby **pozorovania**: odpoveď nedošla, je chybná, zle zaznamenaná, neúplná, nepoužiteľná, ...
- **náhodné** chyby: spočívajúce v tom, že získaní výsledky sa vzťahujú (extrapolujú) na celý základný súbor.

Základné pojmy:

Projekt výberového skúmania musí mať najprv stanovené ciele a spôsob využitia výsledkov.

Uvedieme definície základných pojmov:

Základný súbor	Množina (konečná alebo nekonečná) objektov (prvkov), ktorých vlastnosti budeme zisťovať a odhadovať pomocou výberu. (Population, resp. Target population) <i>Základný súbor je súbor všetkých štatistických jednotiek, ktoré spĺňajú dané kritériá. Napr. dospelá populácia SR, populácia SR vo veku 15 až 64 rokov.</i>
Štatistická jednotka	Prvok základného súboru. (Statistical Unit) <i>Štatistické jednotky musia byť jednoznačne definované – napr. osoby štátny príslušníci SR, účastníci EKOMSTAT' 2007</i>
Opora výberu	Úplný zoznam (najlepšie v elektronickej forme) všetkých objektov základného súboru. (Sample frame)
Výber	Podmnožina základného súboru. (Sample)

Reprezentatívny výber	Podmnožina základného súboru vybraná tak aby reprezentovala záujmový základný súbor, teda aby výberový súbor mal rovnaké zastúpenie. (Representative sample)
Výberová jednotka	Prvok výberového súboru. (Sampling Unit) <i>Niekedy sa tak označuje aj prvok opory výberu – jednotka na vyberanie. Prvky základného súboru, ktoré budeme vyberať musia byť jednoznačne definované. Napr. pri súbore osôb môže byť jednotkou jednotlivá osoba, domácnosť, byt, mesto. V poľnohospodárstve podnik, pozemok. Pri kontrole kvality výrobok, alebo jeho časť. Ak jednotka zahŕňa viac prvkov súboru musí každý prvok patriť iba do jednej zloženej jednotky.</i>
Metóda konštrukcie výberového súboru	Náhodný – pravdepodobnostný. Nenáhodný – empirický (zámerný výber).
Pravdepodobnostný výber	Výberová metóda, ktorá určuje špecifickú pravdepodobnosť vybratia pre prvky základného súboru. (Probability sample) <i>Pri homogénnych prvkoch základného súboru sú tieto pravdepodobnosti spravidla rovnaké.</i>
Nepravdepodobnostný výber	Metóda výberu bez špecifikácie pravdepodobnosti výberu. (Non-probability sample)

Pre realizáciu výberu je potrebné:

Pre výber vzorky - určiť:

rozsah výberu; výberový postup; charakteristiku (priemer, sumu, podiel), ktorú budeme počítať a jej teoretické zdôvodnenie.

Metódy na stanovenie rozsahu výberu sú v každej učebnici výberových metód napr.: Mace A.E.: Sample-size determination. Reihold. New York 1964.

Spôsob (metóda) zberu: dotazovateľ (face-to face), poštou, telefonicky, počítačovo asistovaný zber, Ak je to možné uprednostňuje sa priamy kontakt (školeného) dotazovateľa s respondentom - chyby zisťovania sú menšie a tiež počet získaných odpovedí je väčší.

Sledovanie priebehu zisťovania - kontrola práce dotazovateľov-anketárov, kontakt s anketármi a pomoc v sporných prípadoch.

Obdobie zisťovania a obdobie zberu dát. Obdobie zisťovania sa vzťahuje k zisťovanej informácii. Obdobie zberu musí zabezpečiť dobrú kontrolu pravdepodobností výberu jednotiek (napr. neutralizáciu sezónnych vplyvov).

Dotazník je dôležitá obsahová stránka a i formálna úprava.

2.2 Základné typy výberov

Keďže príspevok je venovaný metóde kvótového výberu stručne uvedieme kvôli úplnosti aspoň „zoznam“ základných typov výberov.

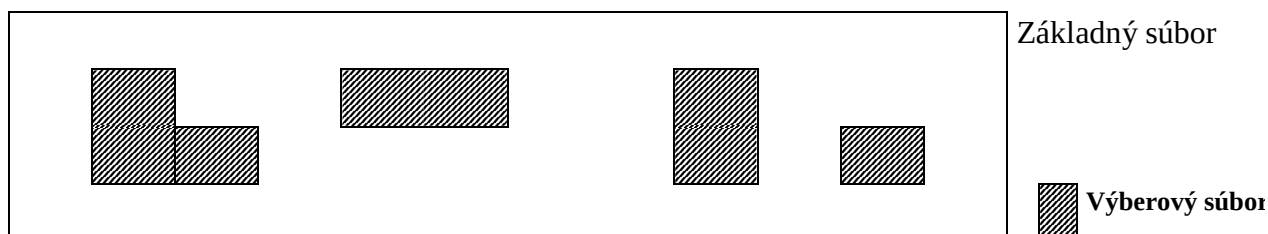
Z nenáhodných (empirických) metód možno spomenúť: kvótový výber, metóda typických jednotiek, itineráre a i. Ich výhodou je jednoduchosť a nízke náklady. Empirické výberové

metódy používame, keď nemáme k dispozícii oporu výberu, ale len niektoré informácie o skúmanom záujmovom základnom súbore.

Z náhodných výberov uvádzame schématické znázornenie, bez bližšej špecifikácie: Jednoduchý náhodný výber, dvojstupňový výber, stratifikovaný náhodný výber a výber skupín.

Schématické znázornenie základných typov náhodných výberov:

Jednoduchý náhodný výber



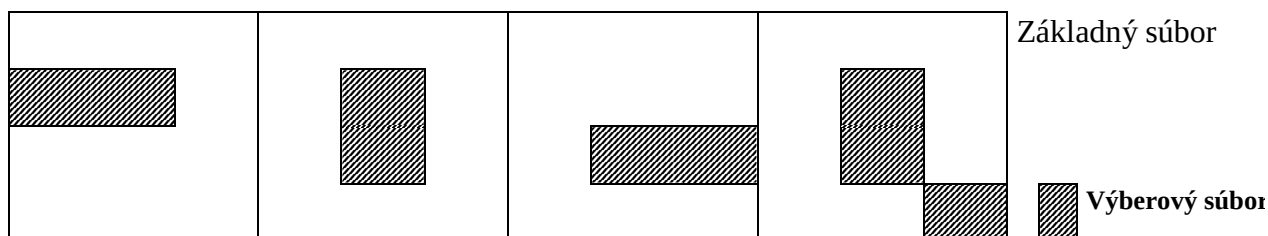
Dvojstupňový výber

Primárne jednotky



Oblasťový výber

Primárne jednotky



Výber skupín

Primárne jednotky



Pri serióznych sociologických (empirických) výskumoch je potrebné, aby výberový súbor bol **reprezentatívny**. **Náhodný výber je reprezentatívny na základe teoretických vlastností**. Pri rozsiahlom základnom súbore však náhode treba pomôcť a preto sa konštruujú zložitejšie typy náhodných výberov. Napr. oblasťový, viacstupňový, výber skupín atď. **Pri iných typoch výberu sa reprezentatívnosť dosahuje rozličnými spôsobmi**. Pri kvótových výberoch reprezentatívnosť výberového súboru (definovanú ako zhodu parametrov (kontrolovaných)

kvótových znakov výberového súboru s parametrami základného súboru) zabezpečujeme rozpisom kvót počtu rozhovorov.

2.2 Odhadovanie základných štatistických charakteristík a výberovej chyby pri kvótovom výbere

Mnohí štatistickí namietajú proti aplikácii vzorcov z teórie náhodných výberov pri kvótových výberoch. V odbornej literatúre sa vyskytujú publikácie, ktoré prinášajú argumenty a aj teoretické zdôvodnenia v prospech aplikácie vzorcov stratifikovaného náhodného výberu na kvótový výber. Z doteraz dostupných zdrojov literatúry uvedieme niekoľko takýchto publikácií. V publikácii **Anděl M., Černý R., Charamza P., Neustadt J.**: Přehled metod odhadu statistické chyby ve výběrových šetřeních. Informační bulletin České statistické společnosti. Číslo 2-3/2004, ročník 15. autori uvádzajú v záverečnej kapitole „V prípade prieskumov face-to-face sa používajú najčastejšie výbery „kvótové“, ktoré aproximujú stratifikované náhodné výbery, kde stratifikácia prebieha väčšinou na regionálnej úrovni, reprezentatívne voči pohlaviu a veku, obvykle aj vzdelaniu. Na výpočet štatistickej chyby možno vtedy odporučiť všeobecné postupy – vid' (48), ale ani aproximácia vzorcami pre jednoduché náhodné výbery z kapitol 2, resp. 4 asi neprinesia hrubšiu chybu.

V niektorých publikáciách sa argumentuje pomocou Bayesovskou teóriou – napr. v:

Gourieroux Christian: Théorie des Sondages. INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES, Economica. 1987.;

Alreck PL, Settle RB.: The Survey Research Handbook. Guidelines and Strategies for Conducting a Survey, Second Edition. Chicago, IL: Irwin Professional Publishing (1995). ISBN: 0-7863-0358-1.

Niektoré empirické štúdie boli realizované na porovnanie kvótového a náhodného výberu. V práci: Grosbras, J.: Statistické metody výběrových šetření. FSÚ Praha 1992 (Preklad z francúzskeho originálu: Méthodes statistiques des sondages) sú uvedené výsledky zisťovania, ktoré že výsledky za meritórne otázky sú prakticky zhodné pre oba typy výberov.

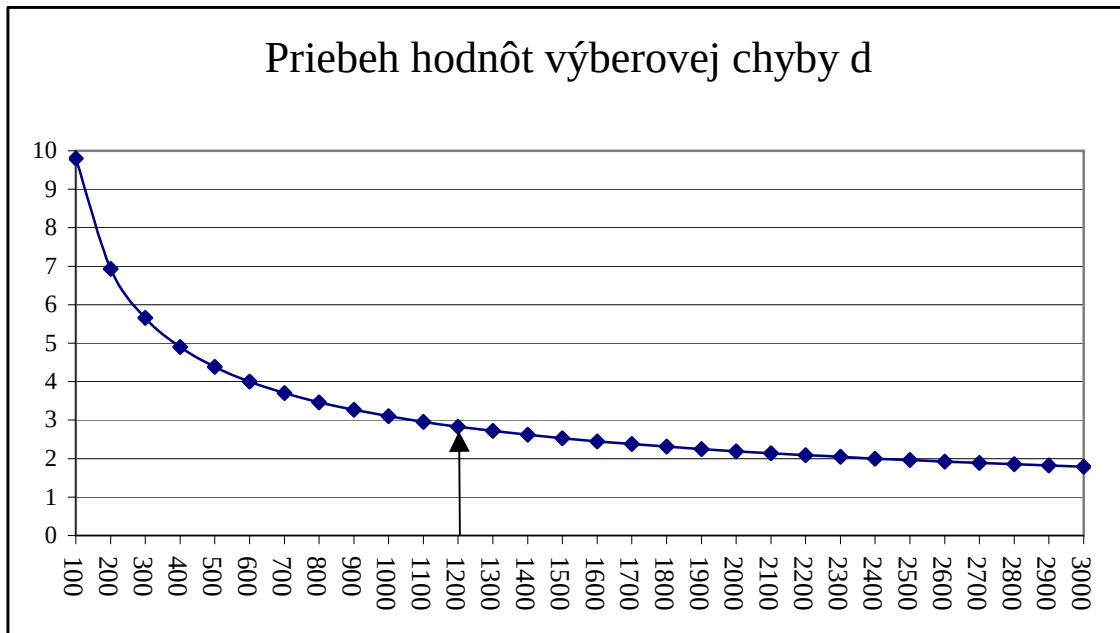
I keď zatiaľ niečo ako priamy dôkaz na aplikáciu vzorcov z náhodných výberov pri kvótovom výbere autor tohoto príspevku nepozná, uznáva uvádzané argumenty a v praxi ich používa.

Uvedené argumenty teda naznačujú možnosť aplikácie vzorcov náhodného výberu, resp. stratifikovaného náhodného výberu pri kvótovom výbere.

Vzorcami na odhad priemeru, podielu a pod. sa nebudeme venovať. Na ilustráciu uvedieme vzorec na výpočet výberovej chyby, ktorý zároveň ilustruje výpočet rozsahu výberu.

Presnosť výsledkov meriame výberovou chybou a s ňou súvisiacim intervalom spoľahlivosti. Výberová chyba, pri dobre realizovanom zbere dát, závisí od rozsahu výberovej vzorky. Problematika merania spoľahlivosti je zložitá. Ilustrujeme ju jednoduchším spôsobom určenia výberovej chyby cez interval spoľahlivosti podielu (percenta) variantu štatistického znaku. Ak uvažujeme koeficient spoľahlivosti 95%, tak interval spoľahlivosti (výberovú chybu) možno aproximačne počítať pomocou vzťahu: $d = \pm 100 \cdot 1,96 \sqrt{p \cdot (1-p) / n}$.

Pri rozsahu výberovej vzorky nad 1000 osôb je výberová chyba menšia než 3%. Priebeh veľkosti výberovej chyby v závislosti od rozsahu výberu je znázornený na grafe.



Pri výskumoch verejnej mienky sa uspokojujeme s výberovou chybou $\pm 3\%$ a preto obvykle volíme rozsah výberu $n=1200$.

3. DEFINÍCIA A ZÁKLADNÉ VLASTNOSTI KVÓTOVÉHO VÝBERU

Kvótový výber považujeme za určitú formu stratifikácie, hoci nespĺňa všetky požiadavky stratifikovaného náhodného výberu.

3.1 Definícia a konštrukcia kvótového výberu

Kvótový výber: vychádza z princípu: *Ak štruktúra výberového súboru za rozhodujúce relevantné znaky zodpovedá ich štruktúre v záujmovom základnom súbore, tak bude výberový súbor reprezentatívny aj za skúmané meritórne znaky.*

Príkladom sú populačné prieskumy: Nech je známe rozdelenie sledovanej populácie (napr. dospelí občania SR) za demografické znaky (pohlavie, vek, vzdelanie, národnosť, veľkosť obce, kraj). *Ak je výberový súbor reprezentatívny podľa vybraných kontrolovaných znakov, tak predpokladáme, že bude reprezentatívny aj pre zisťované meritórne znaky.*

Výberová vzorka sa vypočíta podľa štruktúry kontrolovaných znakov v základnom súbore.

Definíciu kvótového výberu budeme špecifikovať na príklade populačných prieskumov. Výberový súbor obvykle zostavíme metódou voľného kvótového výberu. Táto metóda zabezpečuje reprezentatívnosť marginálne za jednotlivé kontrolované znaky. Teoreticky si možno predstavovať, že navzájom zviazané (krížové) kvóty sú vhodnejšie, ale problém je v raste počtu kombinácií. Ak uvažujeme 6 kontrolovaných znakov s počtom variantov: 2,6,4,3,5,8, tak máme 5760 kombinácií. Ak stanovíme rozsah výberu $n=1200$, tak v krížových kvótach nemôžeme zabezpečiť všetky kombinácie. Voľný kvótový výber má výhody:

- práca s marginálnymi kvótami je rýchlejšia.
- veľmi často nie sú krížové rozdelenia k dispozícii, marginálne však áno.
- experimentálne štúdie ukázali, že krížové kvóty neposkytujú v porovnaní s marginálnymi podstatný prírastok presnosti.

Pri aplikácii metódy kvótového výberu musíme brať do úvahy aj spôsob realizácie zberu dát, napr. telefonický prieskum, face-to-face rozhovory pomocou anketárov. **V ďalšom budeme**

študovať konštrukciu kvótového výberu pre rozhovory pomocou siete anketárov. Najprv musíme poznať alokáciu anketárskej siete.

Pri celoslovenských reprezentatívnych populačných výskumoch je výhodné alokovať anketársku sieť v závislosti od veľkosti obce (počtu jej obyvateľov) podľa okresov, resp. krajov. Podľa obvykle realizovaného rozsahu výskumu stanovíme počet anketárskych pozícií a potom ich rozloženie podľa prvé uvedených demografických charakteristík. Uvedieme postup konštrukcie rozloženia anketárskych pozícií podľa kraja a veľkostnej skupiny obce pre príklad záujmového základného súboru dospelých populácie SR, teda vo veku 18 a viac rokov. Najprv uvedieme bežne používané kvótové znaky: pohlavie (01=muž, 02=žena), vek (01=18-24, 02=25-29, 03=30-39, 04=40-49, 05=50-59, 06=60 a viac rokov), vzdelanie (01=základné, 02=stredné bez maturity, 03=stredoškolské s maturitou, 04=vysokoškolské), národnosť (01=slovenská, 02=maďarská, 03=iná), veľkostná skupina obce (VSO) (01=do 1999 obyv., 02=2000- 9999, 03=10000-49999, 04=50000-9999 05=nad 100000), kraj (01=Bratislavský, 02=Trnavský, 03=Trenčiansky, 04=Nitriansky, 05=Žilinský, 06=Banskobystrický, 07=Prešovský, 08=Košícký).

Postup konštrukcie alokácie anketárskej siete:

Na ilustratívne výpočty boli použité dáta:

Zdroj dát : Vekové zloženie obyvateľstva SR (v roku 2005), Demografia a sociálna štatistika. ŠÚ SR, Bratislava 2006. (za znaky: pohlavie, vek, vso a kraj) , VZPS4q2005 a EU SILC 2005 za znak vzdelanie, SODB 2001 za znak národnosť.

Predpokladajme, že obvyklý rozsah výberovej vzorky je 1200 a počet rozhovorov, ktoré bude realizovať jeden anketár bude rovnaký a rovný 6, potom budeme budovať anketársku sieť s 200 anketárskymi pozíciami.

Tabuľka ilustruje teoretické rozloženie počtu obyvateľov podľa krajov a VSO, napr. dospelých populácie.

Kraj/VSO	VSO1	VSO2	VSO3	VSO4	VSO5	Spolu
Bratislavský	N ₁₁	N ₁₂	N ₁₃	N ₁₄	N ₁₅	N _{1.}
Trnavský	N ₂₁	N ₂₂	N ₂₃	N ₂₄	N ₂₅	N _{2.}
Trenčiansky	N ₃₁	N ₃₂	N ₃₃	N ₃₄	N ₃₅	N _{3.}
Nitriansky	N ₄₁	N ₄₂	N ₄₃	N ₄₄	N ₄₅	N _{4.}
Žilinský	N ₅₁	N ₅₂	N ₅₃	N ₅₄	N ₅₅	N _{5.}
Banskobystrický	N ₆₁	N ₆₂	N ₆₃	N ₆₄	N ₆₅	N _{6.}
Prešovský	N ₇₁	N ₇₂	N ₇₃	N ₇₄	N ₇₅	N _{7.}
Košický	N ₈₁	N ₈₂	N ₈₃	N ₈₄	N ₈₅	N _{8.}
Spolu	N _{.,1}	N _{.,2}	N _{.,3}	N _{.,4}	N _{.,5}	N

Potrebný počet anketárov podľa krajov a VSO získame jednoduchým výpočtom podľa vzorcov:

$P_{ij} = P \cdot N_{ij}/N$, $P_{i.} = P \cdot N_{i.}/N$, $P_{.j} = P \cdot N_{.j}/N$, kde $i=1, 2, \dots, 8$ a $j=1, 2, \dots, 5$, a pre našu konkrétnu sieť je $P=200$.

Ilustračná tabuľka rozloženia počtov anketárov podľa kraja a VSO vyžaduje samozrejme zaokrúhliť čísla na jednotky. To si vyžaduje najprv zaokrúhliť počty anketárov marginálne podľa krajov a podľa VSO. **V realite SR naviac je celkom 8 políčok tabuľky reálne nulových a síce nuly sú pre P_{14} , P_{84} , P_{25} , P_{35} , P_{45} , P_{55} , P_{65} , P_{75} .**

Kraj/VSO	VSO1	VSO2	VSO3	VSO4	VSO5	Spolu
Bratislavský	P ₁₁	P ₁₂	P ₁₃	P ₁₄	P ₁₅	P _{1.}
Trnavský	P ₂₁	P ₂₂	P ₂₃	P ₂₄	P ₂₅	P _{2.}
Trenčiansky	P ₃₁	P ₃₂	P ₃₃	P ₃₄	P ₃₅	P _{3.}
Nitriansky	P ₄₁	P ₄₂	P ₄₃	P ₄₄	P ₄₅	P _{4.}
Žilinský	P ₅₁	P ₅₂	P ₅₃	P ₅₄	P ₅₅	P _{5.}
Banskobystrický	P ₆₁	P ₆₂	P ₆₃	P ₆₄	P ₆₅	P _{6.}
Prešovský	P ₇₁	P ₇₂	P ₇₃	P ₇₄	P ₇₅	P _{7.}
Košický	P ₈₁	P ₈₂	P ₈₃	P ₈₄	P ₈₅	P _{8.}
Spolu	P _{1.}	P _{2.}	P _{3.}	P _{4.}	P _{5.}	P

Vychádzajúc z tejto tabuľky a údajov o počte obyvateľov podľa okresov a VSO rozmiestnime potom anketárske pozície až na úroveň okresov. Poznámka: Vzhľadom na veľký počet okresov v SR (presne 79) je nutné niektoré okresy zlúčiť inak by sme nedokázali stanoviť rozumným spôsobom počty rozhovorov pre danú anketársku pozíciu a pre rozpis podľa ostatných kontrolovaných kvótových znakov. **Konkrétne rozloženie anketárskych pozícií neuvádzame, pretože je to dôverný údaj konkrétnej výskumnej agentúry.**

Pokračujeme stanovením kvót výberovej vzorky **podľa znakov kraj x VSO**. Potup je v tomto prípade už jednoduchý. Stačí vynásobiť počty anketárov v tabuľke počtom rozhovorov pre jedného anketára, čo je v našom prípade číslo 6. Alternatívny postup by bol podobne ako pri výpočte počtu anketárov iba by sme namiesto čísla 200 použili vo vzorcoch rozsah výberu $n=1200$.

Ďalej pokračujeme výpočtom marginálnych kvót za jednotlivé kontrolované znaky a síce: pohlavie, vek, národnosť a vzdelanie, potom výpočtom kvót podľa kraja a jednotlivých kontrolovaných znakov.

Na ilustráciu uvedieme marginálne kvóty pre znak pohlavie a podľa kraja a pohlavia. Okrem príslušného počtu obyvateľov SR, stav ku 31.12 2005, podľa údajov ŠÚ SR sú v tabuľke už zaokrúhlené príslušné kvóty počtu rozhovorov.

V tabuľke uvádzame zvlášť aj kvóty za mestá Bratislava a Košice, aby sme lepšie zabezpečili príslušné kvóty.

Súbor	N18av	muži18av	ženy18av	n18av	n_muži18+	n_ženy18+
SR	4 255 128	2 035 191	2 219 937	1200	574	626
Bratislavský	500 719	233 188	267 531	144	67	77
Trnavský	445 811	214 251	231 560	126	61	65
Trenčiansky	483 428	233 515	249 913	138	67	71
Nitriansky	572 854	272 694	300 160	162	77	85
Žilinský	538 204	260 719	277 485	150	73	77
Banskobystrický	523 511	248 525	274 986	150	71	79
Prešovský	596 394	288 913	307 481	162	78	84
Košický	594 207	283 386	310 821	168	80	88
Bratislava	358005	164734	193271	108	50	58
Košice	187769	87817	99952	54	25	29

Pri zaokrúhľovaní „do kríža“ najmä pri znakoch s väčším počtom kategórií, napr. vek sa prejaví „čaro“ takéhoto zaokrúhľovania. Je jasné, že v takomto prípade môže vzniknúť viac riešení, čo ale skúsenému „tvorcovi“ výberovej vzorky nespôsobí problémy. Ilustovaný postup dáva tiež dobrý základ pri konštrukcii rozpisu počtu rozhovorov na anketársku sieť ako uvidíme neskôr.

3.2 Konštrukcia rozpisu počtu rozhovorov na anketársku sieť

Uvažujeme anketársku sieť s 200 pozíciami. Rozloženie podľa krajov a VSO sme uviedli v predošlej kapitole. Ďalej je potrebné urobiť detailnejší rozklad anketárskych pozícií podľa okresov (resp. spojených okresov) a VSO. **Tým explicitne vytvárame rozloženie anketárov, ktoré pri dobrej návratnosti zabezpečuje reprezentatívnosť podľa znakov kraj a VSO.** Dáta ŠÚ SR či už zo sčítaní, alebo bilancie pohybu obyvateľstva umožňujú vytvoriť podklad na výpočet kvót za ostatné kontrolované znaky: pohlavie, vek, národnosť a vzdelanie podľa okresov x VSO, čo vytvára dobrý základ na konštrukciu kvót na jednotlivé anketárske pozície. Treba však poznamenať, že dáta za vzdelanie a národnosť sú žiaľ k dispozícii iba za sčítania a tak v medziobdobí medzi dvomi sčítaniami, čo predstavuje 10 rokov je potrebné tieto dáta odhadovať. Štruktúra znaku národnosť sa nemení tak rýchlo a preto ju obvykle aktualizujeme iba po sčítaní, čo bolo naposledy v roku 2001. Na odhad štruktúry znaku vzdelanie sme využili okrem dát zo sčítania dva prieskumy ŠÚ SR : VZPS4q2005 a EU SILC 2005.

Naším cieľom je skonštruovať matice ktoré majú 200 riadkov pre jednotlivé anketárske pozície a príslušný počet stĺpcov za znaky pohlavie, vek, národnosť a vzdelanie. V príklade ktorý uvádzame sme zvolili rozsah výberu $n=1200$ a teda celkový počet rozhovorov pre jedného anketára je 6. Z dát o počte obyvateľov sledovaného záumového základného súboru máme matice z riadkov podľa okresov x VSO pre znaky pohlavie, vek, národnosť a vzdelanie. Keďže počet riadkov v týchto maticiach nedosahuje číslo 200, rozpíšeme ich do 200 anketárskych pozícií a prepočítame na počet rozhovorov = 6. Tak dostávame základ pre rozpis počtu rozhovorov na anketársku sieť. Nasleduje zaokrúhľovanie. Využijeme pritom prv vypočítané kvóty výberovej vzorky za uvedené znaky podľa krajov a marginálne za celú SR. Zaokrúhľujeme tak, aby pri každom znaku bol súčet počtu rozhovorov pre každú anketársku pozíciu bo rovnaký a v konkrétnom príklade rovný číslu 6, zároveň musí byť súčet počtu rozhovorov v danom stĺpci pre variant znaku napr. muži rovný príslušnej kvóte za kraj a tiež kvóte celkom. Tu je samozrejme viac možných riešení, čo nám umožňuje rôznymi spôsobmi obmieňať konkrétne kvóty pre daného anketára a zabráňovať mu v tvorení panelov.

Ilustrujme výpočet za znak národnosť a jeden kraj:

ank okr	názov okr	VSO n1200	slov	mad	ina
25 201	Dunajská Streda	1	6	5	0 1
26 201	Dunajská Streda	1	6	5	0 1
27 201	Dunajská Streda	2	6	5	0 2
28 201	Dunajská Streda	3	6	5	0 1
29 202	Galanta	1	6	4	0 2
30 202	Galanta	2	6	4	0 1
31 202	Galanta	3	6	3	0 2
32 202	Galanta	3	6	0	0 1
33 203,2	Hlohovec, Piešťany	1	6	0	0 2
34 203,2	Hlohovec, Piešťany	1	6	0	0 2
35 203,2	Hlohovec, Piešťany	2	6	0	0 1
36 203,2	Hlohovec, Piešťany	3	6	0	1 1
37 205,21	Senica, Skalica	1	6	0	0 1

38 205,21	Senica, Skalica	2	6	0	0	1
39 205,21	Senica, Skalica	3	6	0	1	2
40 205,21	Senica, Skalica	3	6	0	0	1
41 207	Trnava	1	6	0	0	2
42 207	Trnava	2	6	0	0	1
43 207	Trnava	4	6	0	0	1
44 207	Trnava	4	6	0	0	1
45 207	Trnava	4	6	0	0	1
	Spolu	126	93	31	2	

Nebudeme prezrádzať celý rozpis, preto na ilustráciu uvádzame malý „výrez“ rozpisu:

Kód_ank	Kód_okr	Názov_okresu	vso	n1200	muži	ženy	1824	2529	3039	4049	5059	60pl	slov	mad	ina	zákl	ucn	strsm	vys
1	101	Bratislava I	5	6	3	3	0	1	1	1	2	1	6	0	0	0	1	3	2
2	101	Bratislava I	5	6	3	3	1	0	1	1	2	1	6	0	0	0	2	3	1
3	102	Bratislava II	5	6	3	3	1	1	1	1	1	1	5	0	1	1	1	2	2
8	103	Bratislava III	5	6	3	3	1	1	1	1	1	1	5	1	0	1	1	2	2
9	103	Bratislava III	5	6	2	4	0	1	1	1	1	2	6	0	0	0	1	3	2
10	104	Bratislava IV	5	6	3	3	1	1	1	1	1	1	5	0	1	1	1	2	2
16	105	Bratislava V	5	6	3	3	1	1	1	1	1	1	6	0	0	1	1	2	2
17	105	Bratislava V	5	6	3	3	1	1	1	1	1	1	6	0	0	1	0	3	2
19	106,107,108	Malacky, Pezinok, Senec	1	6	2	4	1	1	1	1	1	1	5	1	0	1	2	3	0
21	106,107,108	Malacky, Pezinok, Senec	2	6	3	3	1	0	2	1	1	1	6	0	0	2	2	2	0
24	106,107,108	Malacky, Pezinok, Senec	3	6	3	3	1	0	1	2	1	1	5	0	1	1	2	2	1
25	201	Dunajská Streda	1	6	2	4	0	1	1	2	1	1	1	5	0	1	2	2	1
32	202	Galanta	3	6	3	3	1	0	2	1	1	1	6	0	0	1	2	3	0
33	203,2	Hlohovec, Piešťany	1	6	3	3	0	1	1	1	1	2	6	0	0	2	2	2	0
34	203,2	Hlohovec, Piešťany	1	6	3	3	1	0	2	1	1	1	6	0	0	2	2	2	0

3.3 Základné vlastnosti kvótového výberu

Kvótový výber je metóda stratifikovaného výberu, kde výber jednotiek v rámci oblastí nie je úplne náhodný, pretože výber respondentov vykonáva anketár.

Základnou vlastnosťou kvótového výberu je **zabezpečenie reprezentatívnosti** podľa kontrolovaných kvótových znakov. Ako už bolo spomenuté vychádza sa z predpokladu, že skúmané meritórne znaky korelujú s kontrolovanými kvótovými znakmi (ktorými bývajú v populačných prieskumoch vhodné demografické znaky) a tým je zabezpečená aj reprezentatívnosť meritórnych znakov.

Hoci štatistickí teoretici skôr odmietajú kvótový výber, mnohí ho obhajujú a tvrdia, že za predpokladu dodržania všetkých zásad pri realizácii zberu dát kvóovým výberom možno získať reliabilné výsledky.

V literatúre sa stretávame s hodnotením výhod a nevýhod kvótového výberu:

Hlavné argumenty proti kvóтовému výberu:

1. Nie je možné odhadovať výberovú chybu, pretože chýba znáhodnenie. Niektorí autori tvrdia, že náhodná chyba je veľmi malá v porovnaní s inými chybami a vychýleniami tohoto výberu. Tento argument sa už pokúšajú vyvrátiť ako sme už spomenuli napr. v práci: Gourieroux Christian: Théorie des Sondages. INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES, Economica. 1987.

2. Anketári môžu zlyhať pri zabezpečení reprezentatívnosti výberu respondentov, napríklad určité skupiny obyvateľov sú „dostupné“ iba v určitých hodinách. Alebo anketári nesprávne vyberú respondentov, napríklad v kategórii respondentov 65 a viacročných ja väčšina zoskupená v rokoch 65 a 66.

3. Kontrola kvót je ponechaná na uvážení anketára.

4. Striktná kontrola terénnej práce anketárov je ťažšia, anketári môžu zaradiť respondentov do kategórií, ktoré potrebujú obsadiť a nie kde reálne patria.

5. Dôležitým faktorom je tiež základný predpoklad metódy aby meritórne znaky korelovali s kontrolovanými kvótovými znakmi.

Poznanie týchto nedostatkov pomáha však výskumnej agentúre pripraviť sofistikovanejšie pokyny na prácu v teréne ako aj ich kontrolu a tým do značnej miery eliminovať ich vplyv.

Hlavné argumenty za kvótový výber:

1. Je menej nákladný. V priemere môžu byť náklady polovičné až tretinové oproti náhodnému výberu.

2. Je jednoduchšie administrovať a menej prácny.

3. Môžu byť realizované rýchlo a tým môžu redukovať „pamäťové“ chyby.

4. Nemusíme mať oporu výberu, potrebujeme kvóty.

3.3 Overovanie reprezentatívnosti kvótového výberu

Potrebujeme overiť zhodu štruktúry kontrolovaných znakov v základnom a výberovom súbore. Uvedme najprv formalizáciu definície reprezentatívnosti a testy reprezentatívnosti z práce : Luha J.: Overovanie reprezentatívnosti výberového súboru. FORUM STATISTICUM SLOVACUM 5/2006. ISSN 1336-7420. SŠDS Bratislava 2006.

Formalizácia definície reprezentatívnosti:

Uvažujme kvalitatívny znak Z s r hodnotami, ktoré okódujeme číslami 1, 2, ..., r .

Označme rozdelenie pravdepodobnosti tohoto znaku v skúmanom základnom súbore:

$$\Pi = (\pi_1, \pi_2, \dots, \pi_r), \text{ kde } \sum_{i=1}^r \pi_i = 1$$

a rozdelenie pravdepodobnosti daného znaku vo výberovom súbore (empirické relatívne početnosti a absolútne početnosti):

$$P = (p_1, p_2, \dots, p_r), \text{ kde } \sum_{i=1}^r p_i = 1 \text{ a } N = (n_1, n_2, \dots, n_r), \text{ pričom } \sum_{i=1}^r n_i = n.$$

V prípade reprezentatívneho výberu platí hypotéza H_0 o zhode uvedených rozdelení pravdepodobnosti. Hypotézu môžeme formulovať nasledovne:

H_0 : Rozdelenie N je multinomické $M(\pi, n)$.

Obvykle stanovíme alternatívnu hypotézu H_1 : neplatí H_0 .

Reprezentatívnosť v prípade viac znakov môžeme uvažovať za každý znak zvlášť. Výberový súbor potom považujeme za reprezentatívny, ak je reprezentatívny za všetky kontrolované znaky.

Testy reprezentatívnosti pre kvalitatívne znaky

CHí-kvadrát test dobrej zhody

Na overenie hypotézy H_0 používame testy dobrej zhody. Testovacia štatistika CHÍ-kvadrát testu dobrej zhody je:

$$X^2 = \sum_{i=1}^r (n_i - e_i)^2 / e_i, \text{ kde } e_i = n \cdot \pi_i, i=1, 2, \dots, r.$$

Táto štatistika má asymptoticky X^2 rozdelenie s $r - 1$ stupňami voľnosti.

Podmienkou použitia testu je splnenie podmienok aproximovateľnosti. Tieto sú známe z teoretických štúdií a praxe. Ak platí nulová hypotéza a rozsah výberu rastie, tak rozdelenie štatistiky X^2 konverguje ku chi-kvadrát rozdeleniu s $r - 1$ stupňami voľnosti, pritom však musí aspoň 20% teoretických početností e_i rovných, alebo väčších ako 5 a žiadna z nich nemá byť menšia ako 1. Pre menšie rozsahy výberu sa odporúča aplikovať exaktný test.

Exaktný test (SPSS)

SPSS generuje exaktné multinomické rozdelenie $M(\pi, n)$ za platnosti nulovej hypotézy a potom počíta exaktnú p-hodnotu testu. Tento test je teda vhodnejší. Pre veľmi rozsiahle výberové súbory postačuje Monte Carlo odhad p-hodnoty testu.

Poznámka:

Pre numerické premenné možno na overenie reprezentatívnosti využiť Kolmogorov-Smirnovov test dobrej zhody alebo hore uvedené testy po vhodnej kategorizácii numerickej premennej.

Testy reprezentatívnosti pre kvalitatívne znaky s dvomi hodnotami:

Špecifickým prípadom sú znaky s dvomi možnými hodnotami. Ako príklad môžeme uviesť pohlavie, alebo znak zisťujúci prítomnosť určitej vlastnosti. Kvôli jednoduchosti budeme označovať hodnoty skúmaného znaku číslami 0 a 1.

V tomto prípade sa hypotéza zhody distribúcií vyjadruje vzťahom:

$H_0: p = \pi$, kde p je podiel vo výberovom súbore a π je podiel v základnom súbore.

Ako alternatívu možno stanoviť obojstrannú alternatívu $H_1: p \neq \pi$, alebo jednostranné alternatívy $H_1: p < \pi$, resp. $H_1: p > \pi$.

Na overenie danej hypotézy, pri výbere o rozsahu n možno použiť:

Binomický test

Za predpokladu platnosti H_0 počítame príslušnú exaktnú p-hodnotu.

Za predpokladu platnosti nulovej hypotézy má náhodná veličina p binomické rozdelenie $B(n, \pi)$.

Nech $p = i/n$, potom pre pravdepodobnosť možných výsledkov platí:

$$P(i) = (n! / ((n-i)! \cdot i!)) \pi^i (1 - \pi)^{n-i}, \text{ pre } i=0, 1, 2, \dots, n.$$

a exaktné p-hodnoty počítame podľa vzorcov:

$$p1 = \begin{cases} \sum_{j=0}^i P(j), & \text{ak } i \leq n \cdot \pi & (\text{alternatíva: } p < \pi) \\ \sum_{j=i}^n P(j), & \text{ak } i > n \cdot \pi & (\text{alternatíva: } p > \pi) \end{cases}, \text{ pre jednostranné alternatívy}$$

$$p2 = \sum_{j \in S(i)} P(j), \text{ kde } S(i) = \{j : P(j) \leq P(i)\}, \text{ pre obojstrannú alternatívu.}$$

Vyhodnotenie testu: Zvolíme si úroveň významnosti α ,

potom ak príslušná p-hodnota $< \alpha$, hypotézu o zhode zamietame a v prípade keď,

ak p-hodnota $> \alpha$, tak hypotézu o zhode môžeme na danej úrovni významnosti prijať a výsledok považovať za reprezentatívny.

z-test

No overenie zhody podielu znaku vo výberovom a základnom súbore možno použiť aj z-test (ak sú splnené podmienky aproximácie, ako sú uvedené v časti 6 a Tabuľke 1).

$$z = (p - \pi) / \sqrt{(\pi(1 - \pi)/n)}.$$

Táto štatistika má asymptoticky $N(0,1)$ rozdelenie.

V tomto prípade môžeme využiť kritické hodnoty štandardizovaného normálneho rozdelenia pre uvedené alternatívne hypotézy.

Pre $H_1: p < \pi$ je kritická oblasť daná nerovnosťou $z < z_{\alpha}$.

Pre $H_1: p > \pi$ je kritická oblasť daná nerovnosťou $z > z_{1-\alpha}$.

Pre $H_1: p \neq \pi$ je kritická oblasť daná nerovnosťou $|z| > z_{1-\alpha/2}$. Kde $|z|$ značí absolútnu hodnotu z .

Príklad: Pre $\alpha = 0,05$, platí $z_{1-\alpha} = 1,64$, $z_{\alpha} = -1,64$ a $z_{1-\alpha/2} = 1,96$.

Vyhodnotenie testu: Uvedieme na príklade pre $\alpha = 0,05$. Ak hodnota testovej štatistiky z nie je v kritickej oblasti, tak nulovú hypotézu môžeme prijať a môžeme tvrdiť že výberový súbor je reprezentatívny podľa daného znaku.

Teda výberový súbor považujeme za reprezentatívny:

pre $H_1: p < \pi$, ak $z > -1,64$,

pre $H_1: p > \pi$, ak $z < 1,64$ a

pre $H_1: p \neq \pi$ ak $|z| < 1,96$.

Príklad SPSS procedúr testovania reprezentatívnosti súboru 18+:

****testovanie reprezentatívnosti výberového súboru****18+*stav ku 31_12_2005, vzdelanie upravené podľa VZPS a EU SILC 2005, narodnosť ku 6_5_2001**SODB2001.

* pohlavie .

NPAR TEST

/CHISQUARE=pohl

/EXPECTED=47.83 52.17

/MISSING ANALYSIS

/METHOD=EXACT TIMER(5).

* vek .

NPAR TEST

/CHISQUARE=vek18av

/EXPECTED=14.47 11.08 18.59 18.40 17.00 20.46

/MISSING ANALYSIS

/METHOD=EXACT TIMER(5).

* národnosť .

NPAR TEST

/CHISQUARE=narod

/EXPECTED=86.14 10.42 3.44

/MISSING ANALYSIS

/METHOD=EXACT TIMER(5).

* vzdelanie .

NPAR TEST

/CHISQUARE=vzdel

/EXPECTED=20.50 28.50 39.00 12.00

/MISSING ANALYSIS

/METHOD=EXACT TIMER(5).

* vekostná skupina obce .

NPAR TEST

/CHISQUARE=VSO

/EXPECTED=30.08 20.85 24.16 12.09 12.82

/MISSING ANALYSIS

/METHOD=EXACT TIMER(5).

* kraj .

NPAR TEST

/CHISQUARE=kraj

/EXPECTED=11.77 10.48 11.36 13.46 12.65 12.30 14.02 13.96
 /MISSING ANALYSIS
 /METHOD=EXACT TIMER(5).

4. REALIZÁCIA ZBERU DÁT POMOCOU KVÓTOVÉHO VÝBERU

Stručne načrtujeme techniku zberu dát pri face-to-face rozhovoroch pomocou anketárov.

Anketári okrem dotazníkov, pokynov ku zberu dát, kde majú taktiež stanovené obdobie terénnej fázy obdržia zozpis počtu rozhovorov.

Ako sme už uviedli, anketári sú alokovaní po celom území SR tak, aby bola implicitne zabezpečená reprezentatívnosť za znaky veľkostná skupina obce a kraj. Počty rozhovorov podľa ostatných kontrolovaných znakov, určené tak aby zabezpečili reprezentatívnosť, sú na základe rozpisu oznámené anketárom v Smernici k danej výskumnej úlohe.

Uvádzame ilustratívny **Príklad smernice pre anketára:**

Anketár: XXX YYYY

Číslo anketára: 12345

S M E R N I C E

pre výskum č.E02/2002

Podľa všeobecných pokynov uvedených v Príručke pre anketára, vyhľadajte

6 osôb tak, aby medzi nimi bolo:

	Počet	Miesto pre čiarky (uskutočnené rozhovory)	Miesto pre krížiky (odmietnutia)
Mužov	3	///.....	+++.....
Žien	3	///.....	++.....

Vo veku

18 až 24 rokov	1	/.....	+
25 až 29 rokov	1	/.....	
30 až 39 rokov	1	/.....	++
40 až 49 rokov	1	/.....	+.....
50 až 59 rokov	1	/.....	
60 rokov a viac	1	/.....	

Národnosti

slovenskej	5		++.....
maďarskej	1		++.....
inej	0		

So vzdelaním

základným	2		++.....
stredoškolským bez maturity	1		+.....
stredoškolským s maturitou	2		+.....
vysokoškolským	1		+.....

Smernica je aj pomôcka pre anketára na evidovanie rozhovorov a odmietnutí.

5. LITERATÚRA

1. Agresti A., Coull Ba.: Approximate is better than "Exact" for interval estimation of binomial proportions. *The American Statistician*. 52:119-126, 1998.
2. Alreck PL, Settle RB.: *The Survey Research Handbook. Guidelines and Strategies for Conducting a Survey*, Second Edition. Chicago, IL: Irwin Professional Publishing (1995). ISBN: 0-7863-0358-1.
3. Anděl M., Černý R., Charamza P., Neustadt J.: Přehled metod odhadu statistické chyby ve výběrových šetřeních. *Informační bulletin České statistické společnosti*. Číslo 2-3/2004, ročník 15.
4. Cochran G.W.: *Sampling Techniques*. Third Edition. Wiles, New York 1977.
5. Crawford I.M.: *Marketing Research and Information Systems*. Kap. 7: Chapter 7: Sampling In Marketing Research:
<http://www.fao.org/docrep/W3241E/w3241e08.htm#quota%20sampling>, FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS Rome, 1997. This publication has previously been issued as ISBN 92-851-1005-3 by the FAO Regional Office for Africa.
6. Čermák V.: *Výběrové statistické zjišťování*. SNTL/ALFA Praha 1980.
7. Čermák V., Vrabec M.: *Teorie výběrových šetření. Část 1. VŠE*, Fakulta informatiky a statistiky Praha 1999.
8. Čermák V., Vrabec M.: *Teorie výběrových šetření. Část 2. VŠE*, Fakulta informatiky a statistiky Praha 1999.
9. Čermák V., Vrabec M.: *Teorie výběrových šetření. Část 3. VŠE*, Fakulta informatiky a statistiky Praha 1999.
10. Čermák V., Vrabec M.: *Teorie výběrových šetření. Sbírka úloh. VŠE*, Fakulta informatiky a statistiky Praha 2003.
11. Elliott, C. and Ellingworth, D. (1997) 'Assessing the Representativeness of the 1992 British Crime Survey: The Impact of Sampling Error and Response Biases' *Sociological Research Online*, vol. 2, no. 4,
<http://www.socresonline.org.uk/socresonline/2/4/3.html>
12. Goodman A.: *Introduction to Data Collection and Analysis*, kap.: Surveys and sampling a Survey techniques:
<http://www.deakin.edu.au/~agoodman/sci101/index.php>, Deakin University (1995, 2003)
13. Gourieroux Christian: *Théorie des Sondages*. INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES, Economica. 1987.
14. Grosbras, J.: *Statistické metody výběrových šetření*. FSÚ Praha 1992 (Překlad z francúzskeho originálu: Méthodes statistiques des sondages).
15. Hastings R.: *ESOMAR Marketing Research Glossary*.
<http://www.esomar.org/web/show/id=45195>
16. Herzmann J., Novák I., Pecáková I.: *Výzkumy veřejného mínění*. Skriptum VŠE Praha, Fakulta informatiky a statistiky, 1995.
17. Chajdiak, J.-Haluška, J.: *Teória výberového skúmania*. ES VŠE, Bratislava 1983.
18. Chajdiak J.: *Štatistika jednoducho*. Statis Bratislava 2003.

19. Chambers R. L., Skinner C. J., Editor(s): Analysis of Survey Data. Series: Wiley Series in Survey Methodology. Published Online: 30 May 2003. Series Editor(s): Walter A. Shewhart, Samuel S. Wilks. Print ISBN: 9780471899877 Online ISBN: 9780470867204. Copyright © 2003 John Wiley & Sons, Ltd
<http://www3.interscience.wiley.com/cgi-bin/bookhome/104536356/>
20. Kanderová, M. – Úradníček, V.: Aplikovaná štatistika vo finančno-ekonomickej praxi. OZ Financ, Banská Bystrica 2005.
21. Kanderová, M. – Úradníček, V.: Štatistika a pravdepodobnosť pre ekonómov. 1. časť. OZ Financ, Banská Bystrica 2005.
22. Kanderová, M. – Úradníček, V.: Štatistika a pravdepodobnosť pre ekonómov. 2. časť. OZ Financ, Banská Bystrica 2007.
23. Luha J.: Meranie spoľahlivosti výsledkov výskumu verejnej mienky. Zborník príspevkov Využitie štatistických metód v sociálno - ekonomickej praxi, EKOMSTAT'94 29.5.-3.6. 1994 Trenčianske Teplice, SŠDS.
24. Luha J.: Matematickoštatistické aspekty spracovania dotazníkových výskumov. Štatistické metódy vo vedecko-výskumnej práci 2003, SŠDS, Bratislava 2003.
25. Luha J.: Reprezentatívnosť vo výskumoch verejnej mienky. FORUM STATISTICUM SLOVACUM 2/2005. ISSN 1336-7420. SŠDS Bratislava 2005.
26. Luha J.: Overovanie reprezentatívnosti výberového súboru. FORUM STATISTICUM SLOVACUM 5/2006. ISSN 1336-7420. SŠDS Bratislava 2006.
27. Mace A. E.: Sample-size determination. Reihold. New York 1964.
28. Motulsky H.: GraphPad PRISM, Version 4.0 Statistical Guide. Statisstical analyses for laboratory and clinical research. 2005. GraphPad Software, Inc.
<http://www.graphpad.com/>.
29. Noelleová E.: Výskum veřejného mínění. Úvod do metod demoskopie, Svoboda, Praha 1968.
30. Pecáková I.: Statistické aspekty terénních průzkumů I. Skriptum VŠE Praha 1995.
31. Příručky SPSS. Ver. 14.
32. Särndal C-E., Swensson B., Wretman J.: Maodel Assisted Survey Sampling. Springer. New York, Berlin, Heidelberg, London, Paris, Tokyo, Honk Kong, Barcelona, Budapest. 1992.
33. Stankovičová, Iveta: Analýza rozptylu a systém SAS. Forum Statisticum Slovaccum. ISSN 1336-7420. - Roč. 1, č. 2 (2005), s. 125-129
34. Stankovičová I., Vojtková M.: Viacrozmerné štatistické metódy pre ekonómov a manažérov. IURA, Bratislava 2007, v tlači.

Adresa autora:

RNDr. Ján Luha, CSc.

Jan.Luha@statistics.sk

Aplikovaný vývoj žiarobetonu - vplyv zrnitosti a použitej suroviny na fyzikálno-chemické parametre

Bujňáková Zuzana, Rusinko Peter

Abstract

The article deals with the use of statistic methods in applied development of concrete. It describes the use of regulation and correlation diagrams, explicitly analyses correlations between input and output parameters.

1. Úvod.

Od zavedenia štatistického riadenia kvality v roku 1998 sa v spoločnosti vo veľkej miere využívajú regulačné a korelačné diagramy, histogramy, Paretova analýza a QFD.

Všetky údaje sú analyzované štatistickým softwarom QI ANALYST a graficky vyhodnocované. Pri hodnotení sa sleduje spôsobilosť procesu pomocou indexov spôsobilosti C_p a C_{pk} . Spôsobilosť procesu sa sleduje hlavne pri procesoch mletia, miešania a lisovania. V týchto procesoch sú schválenými metódami regulačné individuálne diagramy, v ktorých sú spracovateľom technologického predpisu určené špecifické medze USL, LSL.

Spracované údaje sú vyhodnocované realizačným tímom, odborom Riadenia kvality a Radou kvality, kde sa prijímajú nápravné opatrenia a zároveň sa vykonáva kontrola prijatých opatrení.

2. Teoretická časť.

2.1 Regulačné diagramy.

Diagram obsahuje strednú priamku (CL – Central Line), ktorá je umiestnená v referenčnej hodnote sledovaného znaku kvality (obyčajne dlhodobá stredná hodnota znaku) a dve štatisticky určené regulačné medze, hornú regulačnú medzu (Upper Control Limit – UCL) a dolnú regulačnú medzu (Lower Control Limit – LCL).

Na hodnotenie spôsobilosti sa používajú *indexy spôsobilosti* (Capability Index): C_p , C_{pk} , C_{pk} , C_{pl} , C_{pm} .

$$C_p = \frac{USL - LSL}{6\sigma}$$

Hodnoty indexu C_p sa interpretujú nasledovne:

- $C_p > 2$ – prehnane dobré, čo môže byť spôsobené príliš zhovievavými požiadavkami zákazníka, alebo nevhodným určením regulačných medzí (šírka pásu $USL - LSL > 12\sigma$)
- $C_p > 1,66$ – veľmi dobré dodržiavanie spôsobilosti procesu (šírka pásu $USL - LSL > 10\sigma$)
- $C_p \geq 1,33$ – dobré dodržiavanie spôsobilosti procesu (šírka pásu $USL - LSL > 8\sigma$)
- $C_p = 1$ - šírka pásu $USL - LSL \approx 6\sigma$
- $C_p < 1$ - spôsobilosť procesu sa nedodržiava
- $C_p < 0,66$ - spôsobilosť sa veľmi nedodržiava (šírka pásu $USL - LSL < 4\sigma$)

$$\sigma = \sqrt{\frac{\sum_{i=1}^k \sum_{j=1}^{n_j} (x_{ij} - \bar{x}_i)^2}{\sum_{i=1}^k (n_i - 1)}}$$

Index spôsobilosti C_p vyjadruje schopnosť výrobného procesu dodržiavať predpísanú variabilitu. Porovnáva sa šírka tolerančného pásu (USL-LSL) so šiestimi smerodajnými odchýlkami reálneho výrobného procesu (6σ).

Nedodržanie spôsobilosti znamená, že v súbore nameraných hodnôt je veľmi vysoká variabilita a je potrebné prijať opatrenia na jej zníženie.

2.2 Korelačné diagramy.

Párový korelačný diagram zobrazuje vzájomnú súvislosť (väzbu) dvoch premenných (dvoch vlastností nameraných súčasne na všetkých štatistických jednotkách skúmaného súboru). Pomáha nájsť súvislosť (koreláciu) medzi dvoma znakmi toho istého súboru. Pri analýze je potrebné určiť, či ide o lineárnu (pozitívnu, negatívnu), nelineárnu alebo žiadnu koreláciu. V prípade potvrdenia závislosti môžeme podniknúť kroky k regulácii kľúčových premenných tak, aby bol proces pod kontrolou.

Lineárna korelácia, pri ktorej body v korelačnom diagrame vytvárajú zoskupenie, je dobre aproximovateľná priamkou. Tesnosť lineárneho vzťahu premenných X a Y číselne vyjadruje koeficient korelácie R_{XY} .

$$R_{xy} = \frac{n \cdot \sum_{i=1}^n X_i \cdot Y_i - \left(\sum_{i=1}^n X_i \right) \cdot \left(\sum_{i=1}^n Y_i \right)}{\sqrt{\left[n \cdot \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2 \right] \cdot \left[n \cdot \sum_{i=1}^n Y_i^2 - \left(\sum_{i=1}^n Y_i \right)^2 \right]}}$$

kde X_i a Y_i sú jednotlivé i-te hodnoty parametrov X, Y.

Koeficient korelácie nadobúda hodnoty z intervalu $<-1, 1>$. Čím je absolútna hodnota väčšia, tým je lineárna závislosť tesnejšia. Pri záporných hodnotách ide o nepriamu a pri kladných o priamu závislosť. Stupeň vzájomnej väzby je:

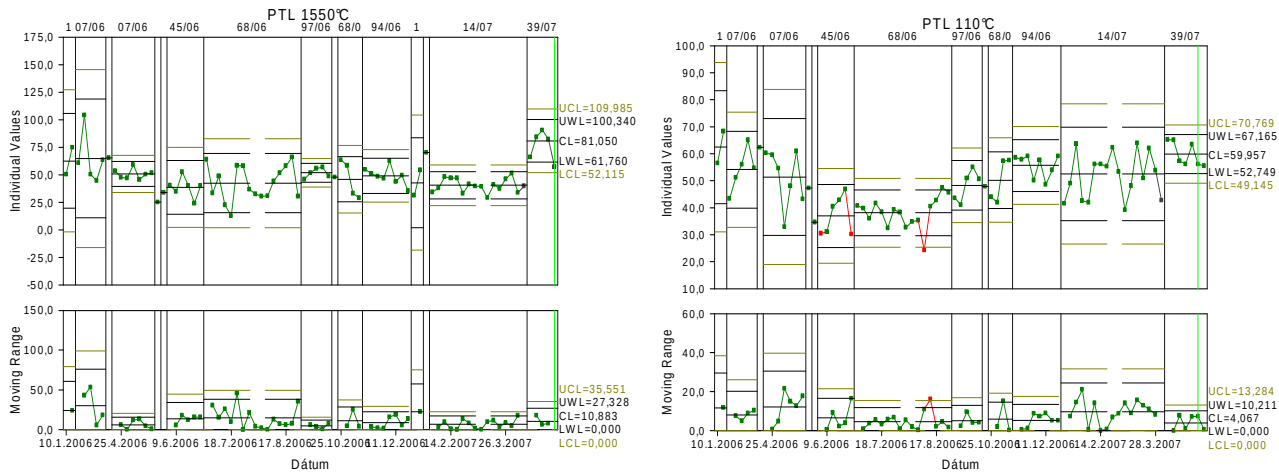
mierny, ak $0,3 \leq |R_{XY}| < 0,5$
významný, ak $0,5 \leq |R_{XY}| < 0,7$
vysoký, ak $0,7 \leq |R_{XY}| < 0,9$
veľmi tesný, ak $0,9 \leq |R_{XY}|$.

3. Praktická časť.

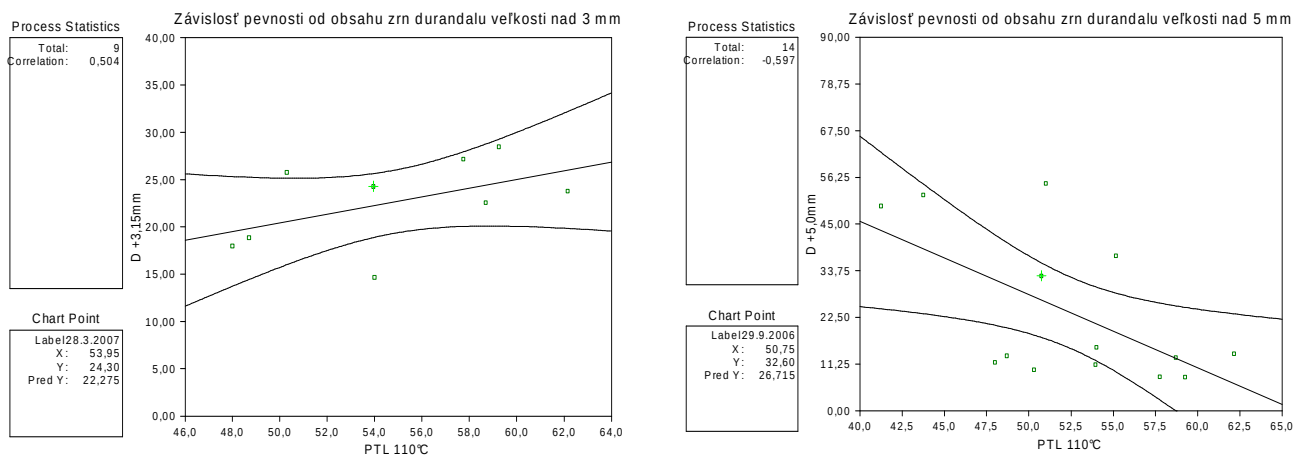
3.1. Aplikovaný vývoj žiarobetónu s novou vstupnou surovinou.

Pri vývoji produktu je nutné mať na pamäti, že špecifikačné medze budú slúžiť ako technické podmienky pre vykonávanie chemicko-fyzikálnych skúšok, ktoré svedčia o kvalite používaného produktu. Je potrebné zhodnotiť technológiu výroby tak, aby bola spôsobilosť procesu čo najvyššia a výskyt chýb čo najnižší. Ak sa rozhoduje o výsledných medziach USL a LSL mali by sa vypočítať zo spôsobilého procesu s indexom $C_p = 1,33$, kedy sa predpokladá, že výroba bude pod štatistickým riadením.

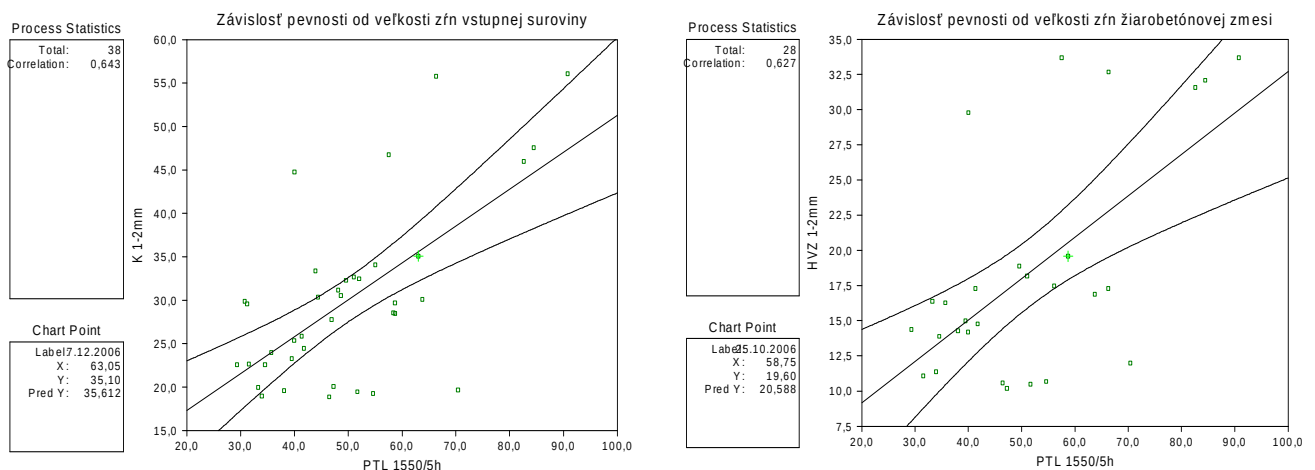
Vývoj výrobku je založený na overení nových vstupných surovín, ktorých vstupné parametre sa mierne odchyľujú od pôvodných požiadaviek. Keďže ide o štandardizáciu existujúceho výrobku, ktorý už trh pozná, sú vopred dané požadované technické parametre a ide teda o meranie a prispôbenie výrobku prevádzkových podmienkam. Pre výpočet σ , USL a LSL boli vzaté do úvahy dva najdôležitejšie parametre pevnosť v tlaku po výpale na 110°C (PTL_{110°C}), a na 1550°C (PTL_{1550°C}). Technické podmienky na daný výrobok garantujú PTL_{110°C} = 40 MPa, PTL_{1550°C} = 40 MPa. Namerané údaje sú graficky vyhodnotené na obr.č.1.



Obr. č. 1 – Namerané hodnoty pevností v tlaku vytriedené podľa receptúr.



Obr. č. 2 - Závislosť pevnosti v tlaku po sušení pri teplote 110°C od obsahu zrn suroviny D



Obr. č. 3 - Závislosť pevnosti v tlaku po výpale pri teplote 1550°C od zrnitosti vstupnej suroviny

Z výsledkov korelácie je zrejmé, ktorým smerom sa vydať pri koncipovaní receptúr. Z obr. č.2 vyplýva, že zrno suroviny D s veľkosťou nad 3mm má mierny vplyv na $PTL_{110^{\circ}C}$ a úmera je priama. Zrno nad 5mm významne vplýva na výslednú pevnosť, avšak úmera je nepriama. Obr. č.3 vykazuje priame závislosti a stupeň väzby významný, medzi pevnosťou v tlaku po výpale a zrnitosťou kameniva a hotovej zmesi s veľkosťou zrna 1-2 mm.

Pri rozhodovaní, ktorú receptúru použiť, definitívne vylučujeme zrno suroviny s veľkosťou nad 5 mm.

Tab.č.1 (vypočítané USL, LSL pre parameter $PTL_{1550^{\circ}C}$)

Číslo receptúry	117/05	07/06	45/06	68/06	97/06	94/06	14/07	39/07
surovina	A 0-4 D 5-8 D 3-5	A 0-4 D 5-8 D 3-5	A 0-4 D 3-5	A 0-4 D 3-5	A 0-4 D 5-8 D 3-5	A 0-4 D 3-5	D 3-5 D 1-3 D 0-1	D 3-5 M 0-4
σ	21,535	8,078	15,769	11,477	4,298	7,754	10,28	9,645
CL	62,85	50,903	40,562	45,035	52,090	49,111	42,243	81,050
USL	148,99	83,215	103,638	90,931	69,182	80,127	83,363	119,63
LSL	-23,29	18,591	-22,514	-0,861	34,989	18,095	1,123	42,47

Tab.č.2 (vypočítané USL, LSL pre parameter $PTL_{110^{\circ}C}$)

Číslo receptúry	117/05	07/06	45/06	68/06	97/06	94/06	14/07	39/07
surovina	A 0-4 D 5-8 D 3-5	A 0-4 D 5-8 D 3-5	A 0-4 D 3-5	A 0-4 D 3-5	A 0-4 D 5-8 D 3-5	A 0-4 D 3-5	D 3-5 D 1-3 D 0-1	D 3-5 M 0-4
σ	10,457	7,165	8,21	3,402	4,608	5,389	9,264	3,604
CL	62,5	54,391	41,525	41,169	48,38	54,883	52,37	59,957
USL	104,328	83,054	74,365	54,777	66,812	76,224	89,476	75,17
LSL	20,672	25,731	9,0	27,561	29,948	33,327	15,314	44,741

V tabuľke č. 1 a 2 sú vypočítané hodnoty σ , CL(priemer), USL, LSL. Podľa nameraných údajov vyhovuje technickým podmienkam receptúra 39/07.

4. Záver.

Žiarobetón sa skladá z kameniva rôznych frakcií a spojiva, používaný v dodanom stave alebo po prídavku jednej alebo viacerých kvapalín.

V receptúre č. 39/07 sa splnila podmienka korelácie z obr. č. 2, čím sa dosiahla minimálna hranica $PTL_{110^{\circ}C}$ 40 MPa a $PTL_{1550^{\circ}C}$ 40 MPa. Keďže vypočítaná LSL pevnosti pri teplote $110^{\circ}C$ je 42,47 MPa a USL 119,63 MPa a rozptyl parametra je ešte stále široký (cca 77 MPa), preto je potrebné sa v ďalšom vývoji zamerať na stabilizáciu parametra. Pri teplote $1550^{\circ}C$ je rozptyl parametra cca 31 MPa.

Keďže korelačná závislosť medzi parametrami vstupov a výstupov je mierna až významná, vzhľadom na použitie viacerých surovín a ich vplyv na výsledné parametre, by bola vhodná viacnásobná regresia.

5. Literatúra.

1. Jozef Chajdiak (2003) : Štatistika jednoducho
2. Michal Tkáč (1998) : QI Analyst 3.0 a štatistické riadenie procesov
3. Jozef Staroň, František Tomšů (2000) : Žiaruvzdorné materiály, výroba, vlastnosti a použitie
4. Jozef Chajdiak (1998) : Štatistické riadenie kvality

Kontakt

Zuzana Bujňáková Ing., Peter Rusinko

REFRAKO

Vstupný areál USS, Košice 045 44

zbujnakova@sk.uss.com

prusinko@sk.uss.com

Analýza meracích systémov (metóda R & R)

Ing. Iveta Ištoňová, Sensus Metering Systems a.s. Stará Turá

This article discusses the R&R method used for evaluation of variations of measuring systems and identification of variability caused by operation or measuring system. The first part of the article mentions theoretical basis of the method and the second part presents specific examples of attribute and variable data.

1. Úvod

Spoločnosť Sensus Metering Systems a.s. Stará Turá vyrába viac ako 1,5 mil. bytových a domových vodomeroch na 15 skúšobných staniciach. Kontrola stability a spôsobilosti merania na týchto staniciach je pre našu spoločnosť rozhodujúca. Ale analýza meracieho systému sa netýka len staníc, ale všetkých procesov, v ktorých sú využívané akékoľvek meracie systémy alebo pri zisťovaní schopností a možností pracovníkov jednoznačne stanoviť porovnávacie parametre výrobkov a procesov. Analýza meracieho systému je v našej spoločnosti používaná:

- ⇒ na validáciu meracieho systému - teda na potvrdenie, že merací systém poskytuje správne hodnoty;
- ⇒ na stanovenie prípadných zdrojov nepresností meracieho systému – človek alebo meradlo.

2. Neistota merania

Koncom 20. – tých rokov minulého storočia bolo pánom Berndtom sformulované tzv. „zlaté“ pravidlo meracej techniky:

„Neistota merania u_C nemá byť väčšia ako $10 \div 20 \%$ z tolerance kontrolovaného rozmeru.“

Toto pravidlo bolo odvodené z predpokladu, že neistota má náhodný charakter s normálnym rozdelením a intervalom spoľahlivosti 95 %. V súčasnosti sú však na procesy bežne kladené požiadavky ako: štatistická zvládnuteľnosť, stabilita procesu, či interval spoľahlivosti 99,73 % ($\pm 3 \sigma$).

3. Merací systém – proces merania

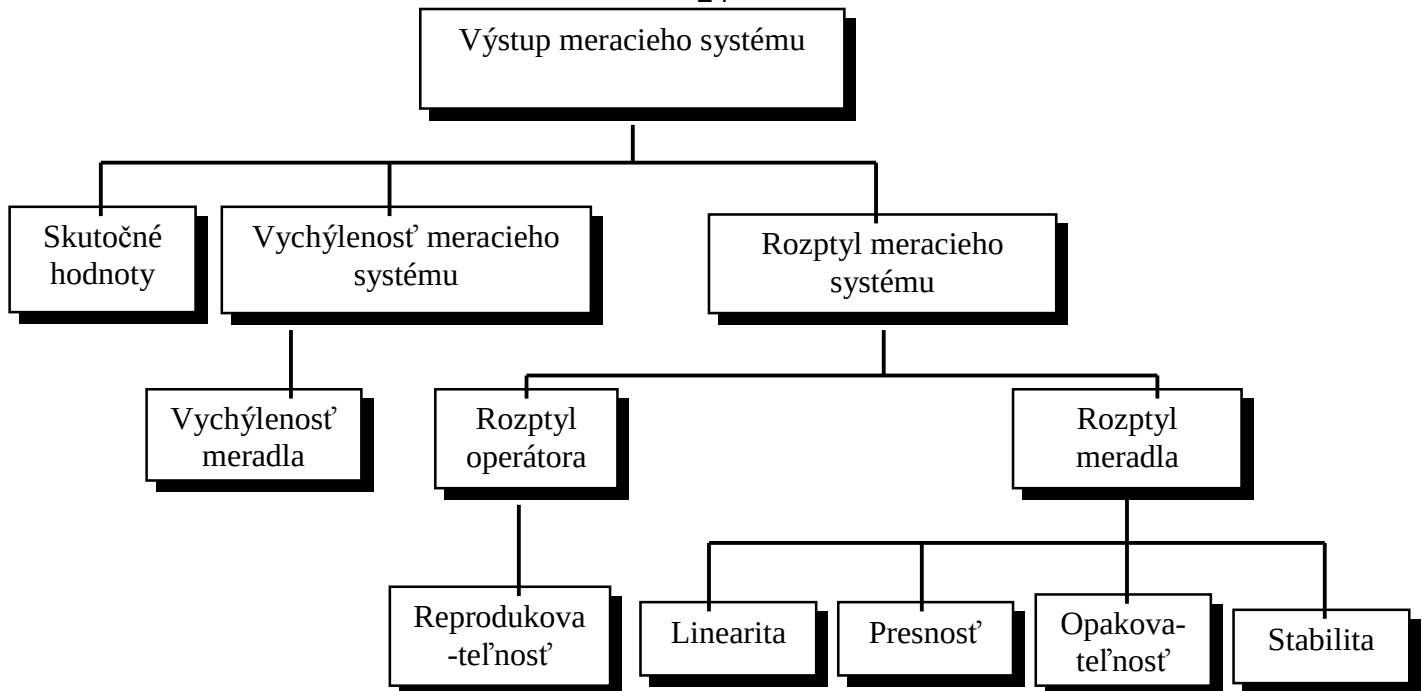
Pri tlaku na trvalé zlepšovanie procesov a znižovanie nákladov na činnosti, ktoré neprinášajú zákazníčkovi úžitok, význam analýzy meracích systémov neustále rastie.

Proces merania by vlastne mohol byť chápaný ako výrobný proces vytvárania údajov, ktoré sú vstupom do procesu analýzy a spracovania nameraných hodnôt a ako taký by mal byť spoľahlivý. Jednou z metód zisťovania spôsobilosti meracieho systému je stanovenie opakovateľnosti a reprodukovateľnosti analyzovaných meracích systémov.

4. Nepresnosti meracieho systému

Všetci uznávame, že namerané hodnoty majú určitý rozptyl. Tento rozptyl ovplyvňuje nielen namerané hodnoty a ich rozloženie a vlastnosti, ale v konečnom dôsledku aj závery analýzy takýchto údajov. Príčiny rozptylu, ktoré sa prejavujú ako chyby systému merania, je možné rozdeliť do nasledujúcich kategórií:

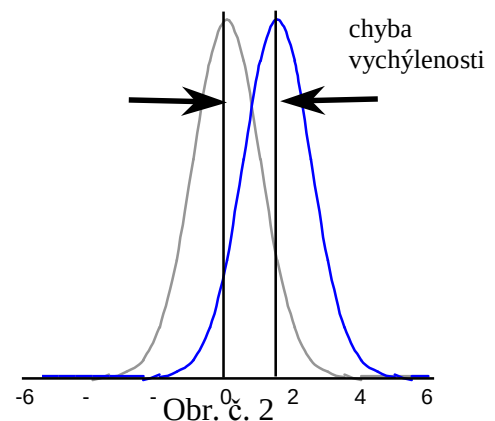
- **správnosť** – tesnosť zhody (rozdiel) medzi priemernou hodnotou výsledkov opakovaných meraní jednej vzorky a pravou (konvenčne pravou) hodnotou meranej veličiny;
- **linearita** – priebeh správnosti meracieho zariadenia v rámci pracovného rozsahu;
- **stabilita** – variabilita priemerných hodnôt výsledkov meraní tých istých vzoriek v rôznom čase;
- **opakovateľnosť** – variabilita merania tej istej vzorky (tých istých vzoriek) pri jej opakovanom meraní jedným operátorom;
- **reprodukovateľnosť** – variabilita priemerných hodnôt výsledkov opakovaných meraní tej istej vzorky (vzoriek) rôznymi operátormi.



Obr. č.1: Nepresnosti meracieho systému

Ako vidno na obrázku č.2 pri analýze nepresností meracieho systému nie je dôležitý len rozptyl nameraných hodnôt, ale aj vychýlenie celého súboru nameraných údajov od požadovanej hodnoty.

Vychýlenie súboru znamená, že hodnoty sú síce správne (ich rozptyl je vyhovujúci), ale celý súbor je posunutý k inej ako požadovanej hodnote. Môžeme teda povedať, že meranie je vlastne proces, ktorý vždy určitým spôsobom skresľuje hodnotenú skutočnosť.

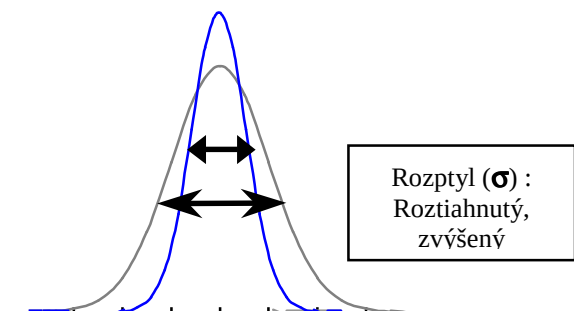


5. Metóda R&R

Zámerom postupu metódy R&R je ohodnotiť rozptyl meracieho systému – stanoviť veľkosť skreslenia skutočnosti. Táto metóda dokáže rozlíšiť variabilitu spôsobenú obsluhou a meracím prístrojom.

$$\sigma_{merania}^2 = \sigma_{skut}^2 + \sigma_{chyba}^2$$

$$\sigma_{chyba}^2 = \sigma_{opakov}^2 + \sigma_{reproduk}^2$$



Obr. č. 3: Rozptyl mer. systému

5.1 R&R metódu je možné použiť pre:

- ⇒ atribútne dáta;
- ⇒ merateľné dáta;

5.2 Podmienky potrebné dodržiavať pri použití metódy R&R:

- najmenej 2 pracovníci;
- najmenej 10 meraných jednotiek (kusov) pre merateľné dáta a 30 meraných jednotiek pre atribútne údaje;
- každá jednotka meraná *náhodne* každým pracovníkom najmenej 2x pri merateľných dátach a min. 3x pri atribútových dátach;
- jednoznačná identifikácia meracích jednotiek (kusov);

5.3 Vyhodnotenie metódy R&R

- a) % R&R < 10 % - merací systém je vyhovujúci;
- b) 10 % < % R&R < 30 % - merací systém je vyhovujúci, ale je nutné ďalej ho skúmať, zlepšiť alebo nahradiť;
- c) % R&R > 30 % - merací systém je nevyhovujúci;

6. Príklad metódy R&R

6.1 Atribútne údaje

Jedným z výrobkov, ktoré ponúkame zákazníčkovi, sú repasované vodomery. Pri ich výrobe sa vyskytol problém – dodržanie dĺžky puzdier vodomeroch – 188 mm. Na meranie tohto rozmeru bol použitý jednoduchý merací prípravok. Pracovník vložil puzdro do prípravku a na základe vlastného uváženia rozhodol, či dĺžka puzdra je vyhovujúca.



6.2 Merateľné údaje

Ako som už spomínala, naša spoločnosť overuje vodomery na 15 skúšobných staniciach. Keďže vodoměr je meradlo určené k platobnému styku, spôsobilosť merania na staniciach je našou prvou prioritou. Ako príklad pre metódu R&R som použila porovnávanie 2 staníc.

6.3 Postup

	Atribútne údaje	Merateľné údaje
1.	Výber vzoriek – 30 puzdier	10 vodomeroch
2.	Označenie vzoriek pre jednoznačnú identifikáciu	
3.	Výber operátorov – r=2; počet opakovaní merania – 3x	Výber staníc – r=2; počet/ poradie opakovaní merania – n = 3; operátori – ten istý prac. tím
4.	Stanovenie experta a definícia stavu atribútu expertom	x
5.	Meranie vzoriek operátormi a zapísanie nameraných údajov do príslušného formulára R&R	
6.	Stanovenie pomernej početnosti n_{ji} – koľkokrát operátor určil taký istý stav atribútu ako expert. i – poradie vzoriek; j – poradie/ počet meraní kde n – počet všetkých meraní 1 operátora Pr.: A- zhoda s expertom; N – nezhoda n = 3*30 = 90 meraní	Výpočet priemerných hodnôt pre jednotlivé skúšky (v rámci stĺpcov): $\bar{x}_{jn} = \frac{\sum_{i=1}^{10} x_i}{N}$; kde N – počet vodomeroch n – poradie skúšky na 1 stanici Pr.: $\bar{x}_{11} = -0,04$
7.	Stanovenie skóre operátora / opakovateľnosti merania (EV): $\bar{n}_j = \frac{\sum_{i=1}^{30} n_{ji}}{n} * 100 \%$ Pr.: $\bar{n}_1 = \frac{89}{3 * 30} * 100 = 99\%$	Vypočíta sa priemer priemerov jednotlivých skúšok pre 1 stanicu: $\bar{x}_j = \frac{\sum_{n=1}^3 \bar{x}_{jn}}{n}$; Pr.: $\bar{x}_2 = 0,09$ kde: n – poradie skúšky na 1 stanici
8.	Určenie priemerného skóre všetkých operátorov: $\bar{n}_{ij} = \frac{\sum_{j=1}^2 n_j}{\sum x_{ij}} * 100 \%$; kde: x_{ij} – počet všetkých meraní operátorov	Výpočet rozpätí meraní 1 vodomera pre 1 stanicu: $R_i = x_{ji \max} - x_{ji \min} $ Pr.: $R_{25} = 0,7 - 0,2 = 0,5$

	Pr.: $\overline{n_{ij}} = \frac{89+88}{(3*30)*2} * 100 = 98,3\%$	
9.	Stanovenie skóre vzoriek: $\overline{x_i} = \frac{n_{ij}}{\sum x_j} * 100 \text{ %}; \text{ kde: } n_{ij} - \text{pomerná}$ početnosť x_j – počet všetkých meraní 1 vzorky Pr.: $\overline{x_{14}} = \frac{2+2}{2*3} * 100 = 66,7\%$	Určenie priemeru rozpätí meraní všetkých vodomeroch na 1 stanici: $\overline{R_j} = \frac{\sum_{i=1}^{10} R_{ij}}{N} \quad \text{Pr.: } \overline{R_2} = 0,42$
10.	Stanovenie reprodukovateľnosti merania (AV): počet prípadov, kde skóre vzorky $\overline{x_j} = 100\%$ / počet všetkých vzoriek „i“ Pr.: $AV = \frac{28}{30} * 100 = 93,3\%$	Určenie rozsahu priemerných rozpätí všetkých vodomeroch na 1 stanici: $R_{\overline{R_j}} = \left \overline{R_{j \max}} - \overline{R_{j \min}} \right ;$ Pr.: $R_{\overline{R_j}} = 0,42 - 0,03 = 0,39$
11.	Určenie hodnoty R&R: $R\&R = 1 - AV = 1 - 0,93 = 0,07 * 100 = 7 \%$	Vypočíta sa priemer rozpätí pre všetky vodomery a všetky stanice: $\overline{R} = \frac{\sum_{j=1}^2 \overline{R_j}}{r}; \quad \text{Pr.: } \overline{R} = \frac{0,03 + 0,42}{2} = 0,225$
12.	$R\&R = 7 \% < 10 \%$ - merací systém je vyhovujúci	Vypočítajú sa priemerné hodnoty vzorky- pre 1 vodomera a všetky stanice: $\overline{x_i} = \frac{\sum_{j=1}^2 x_{ijn}}{\sum_{j=1}^2 n_{ij}};$ Pr.: $\overline{x_1} = \frac{0 + 0 + 0 + 0,2 + 0,2 + 0,8}{6} = 0,2$
13.	xxx	Vypočíta sa rozsah priemerných hodnôt vzoriek: $\overline{R_p} = \left \overline{x_{i \max}} - \overline{x_{i \min}} \right ;$ Pr.: $\overline{R_p} = 0,883 - (-0,7) = 1,583$
14.	xxx	Stanoví sa variabilita / rozptyl vzoriek: $PV = R_p * K_3$ Pr.: $PV = 1,583 * 1,67 = 2,644$
15.	xxx	Stanoví sa opakovateľnosť merania: $EV = \overline{R} * K_1; \quad \text{Pr.: } EV = 0,225 * 3,05 = 0,686$
16.	xxx	Stanoví sa reprodukovateľnosť merania: $AV = \sqrt{\left(\overline{R} * K_2 \right)^2 - \frac{EV^2}{N * r}};$ Pr.: $AV = \sqrt{(0,39 * 3,65)^2 - \frac{0,686^2}{10 * 2}} = 1,415$
17.	xxx	Určenie R&R: $R \& R = \sqrt{EV^2 + AV^2};$ Pr.: $R \& R = \sqrt{0,686^2 + 1,415^2} = 1,573$
18.	xxx	Stanovenie celkového rozptylu: $TV = \sqrt{R \& R^2 + PV^2}$

		Pr.: $TV = \sqrt{1,573^2 + 2,644^2} = 3,077$
19.	xxx	Stanovenie percentuálnych pomerov k celkovému rozptylu: Opakovateľnosť $\%EV = \frac{EV}{TV} * 100 = 22,3\%$ Reprodukovateľnosť $\%AV = \frac{AV}{TV} * 100 = 46\%$ Hodnota R&R $\%R \& R = \frac{R \& R}{TV} * 100 = 51,12\%$ Rozptyl vzoriek $\%PV = \frac{PV}{TV} * 100 = 85,9\%$
20.	xxx	R&R = 46 % > 30 % - merací systém nie je vyhovujúci – parametre nastavenia staníc nie sú rovnaké

7. Záver

Použitie metódy R&R je len prvým krokom , orientáciou v nastavení meracieho systému. Ak je merací systém správne nastavený a vhodný pre zamýšľané použitie, metóda R&R nám o tom dá relatívne rýchlu a jednoduchú odpoveď. Ak je systém na hranici akceptovateľnosti alebo je neakceptovateľný, je nutné pokračovať v skúmaní systému, hľadať zdroje chýb a prijímať opatrenia na ich nápravu.

Ak je zdrojom rozptylu opakovateľnosť, je vhodné zamerať sa napr. na prestavenie alebo opravu zariadenia. Ak je zdrojom rozptylu reprodukovateľnosť, je vhodné zamerať sa napr. na preškolenie alebo výmenu pracovníkov, ktorý sa meracieho procesu zúčastňujú.

Použitá literatúra:

Metodické pokyny spoločnosti Sensus Metering Systems

Kontakt

Ing. Iveta Ištoňová, Sensus Metering Systems a.s. Stará Turá, iveta.istonova@sensus.com

Formulár na analýzu opakovateľnosti a reprodukovateľnosti meracieho systému R&R

Atribútne údaje

Hodnotený údaj: dĺžka puzdra

Dátum: 19.4.2005

Spracoval: xxx

Poradie	Expert	Operátor 1				Operátor 2				Skóre vzorky %
		1. meranie	2. meranie	3. meranie	Pomerná početnosť	1. meranie	2. meranie	3. meranie	Pomerná početnosť	
1	A	A	A	A	3	A	A	A	3	100%
2	N	N	N	N	3	N	N	N	3	100%
3	N	N	N	N	3	N	N	N	3	100%
4	N	N	N	N	3	N	N	N	3	100%
5	N	N	N	N	3	N	N	N	3	100%
6	N	N	N	N	3	N	N	N	3	100%
7	A	A	A	A	3	A	A	A	3	100%
8	N	N	N	N	3	N	N	N	3	100%
9	N	N	N	N	3	N	N	N	3	100%
10	N	N	N	N	3	N	N	N	3	100%
11	N	N	N	N	3	N	N	N	3	100%
12	N	N	N	N	3	N	N	N	3	100%
13	N	N	N	N	3	N	N	N	3	100%
14	N	N	A	N	2	N	A	N	2	67%
15	N	N	N	N	3	N	N	N	3	100%
16	N	N	N	N	3	N	N	N	3	100%
17	N	N	N	N	3	N	N	N	3	100%
18	A	A	A	A	3	A	A	A	3	100%
19	N	N	N	N	3	N	N	N	3	100%
20	N	N	N	N	3	N	N	N	3	100%
21	A	A	A	A	3	A	A	A	3	100%
22	N	N	N	N	3	A	N	N	2	83%
23	N	N	N	N	3	N	N	N	3	100%
24	A	A	A	A	3	A	A	A	3	100%
25	A	A	A	A	3	A	A	A	3	100%
26	N	N	N	N	3	N	N	N	3	100%
27	A	A	A	A	3	A	A	A	3	100%
28	A	A	A	A	3	A	A	A	3	100%
29	N	N	N	N	3	N	N	N	3	100%
30	N	N	N	N	3	N	N	N	3	100%
OPERÁTOR č.1 - skóre-->					99%	OPERÁTOR č.2 - skóre-->				98%

REPRODUKOVATEĽNOSŤ ->

93 %

<=95% Conf Interval=>

96 %

100 %

PRIEMER VŠETKÝCH OPERÁTOROV

98 %

R&R ->

7 %

Hodnotenie založené na tejto štúdii, merací systém
je NEAKCEPTOVATEĽNÝ

Formulár na analýzu opakovateľnosti a reprodukovateľnosti meracieho systému R&R

Merateľné údaje

Hodnotený údaj: porovnanie staníc
Limity:

Dátum: 18.3.2006

Spracoval: xxx

	Stanica 1				Stanica 2				priem.hodn.	
Sample	1.meranie	2.meranie	3.meranie	Rozsah R1	1.meranie	2.meranie	3.meranie	Rozsah R2	vzorky	
1	0	0	0	0,000	0,2	0,2	0,8	0,600	0,200	
2	-0,2	-0,2	-0,2	0,000	-0,3	-0,2	-0,4	0,200	-0,250	
3	0,8	0,9	0,8	0,100	0,9	0,9	1	0,100	0,883	
4	-0,4	-0,5	-0,4	0,100	-0,4	-0,3	-0,3	0,100	-0,383	
5	0,2	0,2	0,2	0,000	0,2	0,2	0,7	0,500	0,283	
6	-0,3	-0,2	-0,2	0,100	-0,4	-0,3	-0,3	0,100	-0,283	
7	0,1	0,1	0,1	0,000	0	0,1	0,7	0,700	0,183	
8	-0,8	-0,8	-0,8	0,000	-0,8	-0,8	-0,2	0,600	-0,700	
9	0,4	0,4	0,4	0,000	0,3	0,4	0,9	0,600	0,467	
10	-0,2	-0,2	-0,2	0,000	-0,2	-0,3	0,4	0,700	-0,117	
priemer $\bar{x}_{jn}$	-0,040	-0,030	-0,030	0,030	-0,050	-0,010	0,330	0,420	Rp= 1,583	
Stanica 1 - priemer xx=>				-0,03333	Sttanica 2 - priemer xx=>				0,09	rozsah priem.
										hodnôt vzoriek

	hodnota	%
OPAKOVATEĽNOSŤ	0,686250	22,31%
REPRODUKOVATEĽNOSŤ	1,415205	46,00%

R&R:	1,572814	51,12%
-----------------	----------	--------

ROZPTYL VZORIEK 2,644167 85,94%

CELKOVÝ ROZPTYL 3,076583

Tolerancia spôsobilosti (+/-)	
Hraničná =>	0,0262136
Akceptovateľná =>	0,0786407

Hodnotenie založené na tejto štúdii, merací systém je
NEAKCEPTOVATEĽNÝ

Posúdenie závislosti nákladov domácností na bývanie od vybraných faktorov metódami regresnej a korelačnej analýzy

Erik Šoltés

Abstract

In our report we analyzed the impact which the selected factors have on households accommodation costs each month. For our analysis we used report called 'Income and Living Conditions Findings' EU SILC 20005, issued by ŠÚSR - Statistical Office of The Slovak Republic. By using methods of regression and correlation analysis we found out which factors influence accommodation costs. We also considered their influence.

Key words: regression and correlation analysis, backward elimination, stepwise regression, all possible regressions, influential observation, multicollinearity, residuals.

1. Úvod

Cieľom príspevku je prostredníctvom metód regresnej a korelačnej analýzy modelovať závislosť nákladov domácností na bývanie od relevantných faktorov. Výsledkom čoho má byť regresný model. Ak tento model bude kvalitný, je možné ho použiť na posúdenie vzťahov medzi nákladmi domácností na bývanie a vybranými premennými ale aj na odhady nákladov na bývanie pri známych hodnotách uvažovaných premenných.

2. Charakteristika uvažovaných premenných

Vychádzali sme z údajov zo *Zisťovania o príjmoch a životných podmienkach* - EU SILC 2005, ktoré vo vybraných domácnostiach v období od 16. mája do 16. júna 2005 zabezpečili krajské správy Štatistického úradu SR. Prostredníctvom metód regresnej a korelačnej analýzy sme modelovali závislosť *celkových mesačných nákladov domácností na bývanie* od vybraných faktorov.

Do regresného modelu ako vysvetľovaná premenná vstupovala premenná *Y - Celkové mesačné náklady domácností na bývanie* (v Sk). Ide o mesačné náklady spojené s právom domácnosti žiť v danom obydľí. Sú v nich zahrnuté aj náklady na služby (voda, elektrina, plyn a kúrenie), ktoré sú výsledkom používania daného obydľia, náklady na pravidelnú údržbu domu/bytu, náklady spojené s poistením domu/bytu a domová (resp. bytová) daň. V prípade vlastníctva rodinného domu i s príľahlým pozemkom je započítaná len daň za rodinný dom bez pozemku (t.j. za zastavanú plochu pozemku).

Uvažovali sme s týmito vysvetľujúcimi premennými:

X1: Nájomné – súčasné nájomné (v Sk) súvisiace s obývaním obydľia.

X2: Vaňa/sprcha - s obmenou 1, ak obydľie je vybavené vaňou alebo sprchou, s obmenou 0, ak obydľie nie je vybavené vaňou alebo sprchou.

X3: Toaleta - s obmenou 1, ak obydľie je vybavené vnútornou splachovacou toaletou, s obmenou 0, ak obydľie nie je vybavené vnútornou splachovacou toaletou.

X4: Nedoplatky-úver - s obmenou 1, ak domácnosť mala v súvislosti s hypotékou alebo platbou nájomného nedoplatky v priebehu posledných 12 mesiacov, s obmenou 0, ak domácnosť nemala takéto nedoplatky.

X5: Nedoplatky-energie - s obmenou 1, ak domácnosť mala v súvislosti s účtami za dodávku energií nedoplatky v priebehu posledných 12 mesiacov, s obmenou 0, ak domácnosť nemala takéto nedoplatky.

X6: Telefón }
 X7: Televízor } Premenné nadobúdajú hodnotu 1, ak domácnosť má (vlastní) zodpovedajúcu položku bez ohľadu na jej počet. V opačnom prípade nadobúdajú hodnotu 0.
 X8: Počítač }
 X9: Práčka }

X10: Obydlie1 - s obmenou 1 pre domácnosti, ktoré obývajú samostatný dom, s obmenou 0 pre ostatné domácnosti.

X11: Obydlie2 - s obmenou 1 pre domácnosti, ktoré obývajú byt v bytovom dome, ktorý má menej ako 10 bytov a s obmenou 0 pre ostatné domácnosti.

X12: Obydlie3 – s obmenou 1 pre domácnosti, ktoré majú iný druh ubytovania, napr. byty v budovách, ktoré sa používajú na iný účel ako na bývanie (materské školy, školy, atď.), s obmenou 0 pre ostatné domácnosti.

Ak všetky tri vyššie uvedené premenné nadobúdajú hodnotu 0, ide o domácnosť, ktorá obýva byt v bytovom dome s najmenej 10 bytmi.

X13: Vlastníctvo – s obmenou 1 pre nájomníka alebo podnájomníka platiaceho bežné nájomné alebo trhovú cenu a s obmenou 0 za ubytovanie prenajímané za zníženú cenu.

X14: Počet izieb – počet izieb k dispozícii pre domácnosť.

X15: Zatekanie strechy – s obmenou 1, ak v obydlí, ktoré obýva domácnosť je zatekajúca strecha, vlhké steny (podlaha, základy) alebo zahŕňajúce okenné rámy, s obmenou 0 pre ostatné obydlia.

X16: Udržiavanie tepla – s obmenou 1, ak obydlie, ktoré obýva domácnosť, má schopnosť udržiavať adekvátne teplo, s obmenou 0 pre ostatné obydlia.

X17: Veľkosť domácnosti – ekvivalentná veľkosť domácnosti (použitá je modifikovaná OECD škála, ktorá dáva váhu 1 pre prvého dospelého, váhu 0,5 pre každého ďalšieho dospelého člena a váhu 0,3 pre každé dieťa mladšie ako 14 rokov).

X18: Urbanizácia-stredná – s obmenou 1, pre domácnosť žijúcu na území s priemerne hustým osídlením a s obmenou 0 pre ostatné domácnosti,

X19: Urbanizácia-hustá – s obmenou 1, pre domácnosť žijúcu na území s hustým osídlením a s obmenou 0 pre ostatné domácnosti.

Ak obidve vyššie uvedené premenné nadobúdajú hodnotu 0, ide o domácnosť, ktorá žije na území s riedkym osídlením.

3. Výber najoptimálnejšej podmnožiny vysvetľujúcich premenných

Zoznam premenných kandidujúcich do regresného modelu je relatívne dlhý. Preto sme na výber, podľa možnosti čo najoptimálnejšej, množiny vysvetľujúcich premenných, použili metódu postupnej eliminácie (*backward elimination*), metódu krokovej regresie (*stepwise regression*) a metódu všetkých možných regresíí (*all possible regression*).

Metóda postupnej eliminácie vychádza z úplného regresného modelu, ktorý obsahuje všetky kandidujúce premenné. V ďalších krokoch sa postupne z modelu vylučujú premenné, ktorých prínos na vysvetlenie variability závislej premennej nie je na zvolenej hladine α štatisticky významný. Spustením procedúry v SAS Enterprise Guide získame výstup s jednotlivými krokmi metódy. Výsledný model je uvedený vo výstupe 1 a zhrňujúca tabuľka vo výstupe 2. Podľa tabuľky analýzy rozptylu vo výstupe 1 je model ako celok štatisticky významný ($p < 0,0001$). V procedúre sme pre vylúčenie premenných zadali hladinu významnosti 0,1. V modeli boli teda ponechané len tie premenné, ktorých prínos, ohodnotený charakteristikou F , bol štatisticky významný na hladine $\alpha = 0,1$. Metódou postupnej eliminácie sa v modeli ponechalo 8 vysvetľujúcich premenných, z ktorých najmenší ($F = 3,42$), ale na hladine významnosti 0,1 ešte štatisticky významný prínos ($p = 0,0650 < \alpha = 0,1$) na vysvetlenie

variability nákladov na bývanie mala premenná *Telefón*. Ostatné premenné boli v jednotlivých krokoch z modelu postupne vylučované, tak ako je uvedené vo výstupe 2.

Výstup 1: Výber regresorov získaný metódou postupnej eliminácie

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	8	4695962488	586995311	341.28	<.0001
Error	418	718942340	1719958		
Corrected Total	426	5414904828			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	-1883.05154	583.86804	17890101	10.40	0.0014
Nájomné	0.93951	0.01950	3992361469	2321.20	<.0001
Počet izieb	677.71664	91.76632	93809643	54.54	<.0001
Veľkosť domácnosti	665.56865	112.38839	60319902	35.07	<.0001
Obydlie1	-1779.82073	428.55773	29665498	17.25	<.0001
Obydlie3	-3312.03613	932.58892	21693389	12.61	0.0004
Vlastníctvo	1849.73103	483.27494	25196858	14.65	0.0001
Telefón	543.30264	293.64276	5887932	3.42	0.0650
Počítač	-289.96844	135.65128	7859076	4.57	0.0331

Výstup 2: Súhrnná tabuľka krokov metódy postupnej eliminácie

Summary of Backward Elimination							
Step	Variable Removed	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	Nedoplatky-energie	18	0.0000	0.8696	18.0470	0.05	0.8284
2	Urbanizácia-hustá	17	0.0000	0.8695	16.1085	0.06	0.8042
3	Práčka	16	0.0000	0.8695	14.1723	0.06	0.8002
4	Obydlie2	15	0.0000	0.8695	12.2331	0.06	0.8046
5	Udržiavanie tepla	14	0.0000	0.8694	10.3787	0.15	0.7017
6	Televízor	13	0.0001	0.8694	8.5760	0.20	0.6553
7	Vaňa/sprcha	12	0.0001	0.8693	6.8424	0.27	0.6037
8	Toaleta	11	0.0001	0.8692	5.1563	0.32	0.5727
9	Zatekanie strechy	10	0.0004	0.8688	4.4444	1.31	0.2531
10	Nedopatky-úver	9	0.0007	0.8681	4.6513	2.24	0.1351
11	Urbanizácia-stredná	8	0.0009	0.8672	5.3052	2.69	0.1018

Metóda krokovej regresie vychádza z modelu, ktorý neobsahuje žiadnu vysvetľujúcu premennú. V jej jednotlivých krokoch sa do regresného modelu začleňujú podstatné vysvetľujúce premenné. Kritériom pre začlenenie vysvetľujúcej premennej je test štatistickej významnosti prínosu, ktorým dodatočne prispieva k -ta vysvetľujúca premenná na vysvetlenie variability závislej premennej. Navyše po každom začlenení niektorej vysvetľujúcej premennej do regresného modelu sa ešte testuje štatistická významnosť regresných koeficientov stojacich pri premenných, ktoré boli do modelu zaradené v predchádzajúcich krokoch. Podľa zhrňujúcej tabuľky vo výstupe 3, mala najväčší vplyv na náklady domácností na bývanie premenná *Nájomné*, ktorá bola do modelu začlenená v 1. kroku. V ďalších krokoch boli podľa veľkosti prínosu (*F Value*) do modelu ešte začlenené premenné *Počet izieb*, *Veľkosť domácnosti*, *Obydlie1*, *Vlastníctvo*, *Obydlie3*, *Počítač*, *Telefón*. Žiadna z nich nemala po začlenení ďalších premenných, štatisticky nevýznamný vplyv na závislú premennú. Preto ani jedna z nich nebola následne v niektorom kroku z modelu vylúčená, a preto stĺpec *Variable Removed* je prázdny. Ako a o koľko sa v jednotlivých krokoch zvyšoval podiel vysvetlenej variability premennej *Celkové mesačné náklady domácností na bývanie* ukazujú stĺpce *Model R-Square* a *Partial R-Square*. Výber regresorov získaný metódou krokovej regresie a metódou postupnej eliminácie je v tomto prípade rovnaký.

Výstup 3: Súhrnná tabuľka metódy krokovej regresie

Summary of Stepwise Selection								
Step	Variable Entered	Variable Removed	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	Nájomné		1	0.7921	0.7921	225.818	1619.01	<.0001
2	Počet izieb		2	0.0502	0.8423	71.1736	134.95	<.0001
3	Veľkosť domácnosti		3	0.0079	0.8502	48.4443	22.38	<.0001
4	Obydlie1		4	0.0065	0.8567	30.0809	19.22	<.0001
5	Vlastníctvo		5	0.0044	0.8611	18.3253	13.36	0.0003
6	Obydlie3		6	0.0039	0.8650	8.1943	12.10	0.0006
7	Počítač		7	0.0011	0.8661	6.6983	3.51	0.0618
8	Telefón		8	0.0011	0.8672	5.3052	3.42	0.0650

Metóda všetkých možných regresíí vyžaduje odhadnúť regresné modely obsahujúce všetky možné kombinácie vysvetľujúcich premenných, resp. regresorov. Pretože uvažujeme 19 kandidujúcich premenných, metódou najmenších štvorcov sa odhadlo $2^{19} = 524\,288$ regresných modelov. V tomto čísle sú zahrnuté: jeden model bez vysvetľujúcej premennej, 19 regresných modelov s jednou vysvetľujúcou premennou, 171 regresných modelov s dvoma vysvetľujúcimi premennými, ..., a nakoniec jeden úplný model (so všetkými 19 vysvetľujúcimi premennými). Na výber najvhodnejšieho modelu, a teda aj na výber najlepšej podmnožiny vysvetľujúcich premenných slúžia najmä tieto kritéria kvality modelu: C_p – Mallowsova štatistika, r_{adj}^2 – upravený (korigovaný) koeficient determinácie, MSE – reziduálny rozptyl, AIC – Akaikeovo informačné kritérium, BIC – Sawaovo Bayesovské informačné kritérium, PC – Amemiyaovo predikčné kritérium, SBC – Swarzoovo Bayesovské kritérium. Keďže všetky zohľadňujú počet vysvetľujúcich premenných v regresnom modeli, môžeme ich použiť aj na porovnanie modelov s rôznym počtom regresorov.

Vo výstupe 4 je výber 10 najlepších modelov podľa kritéria – C_p . Vysoké hodnoty Mallowsvej štatistiky informujú o skreslení modelu. Mallows (1973) navrhol vybrať model s nízkou hodnotou C_p , ktorá sa blíži k počtu jeho parametrov p . Hocking (1976) navyše navrhol z hľadiska predikcie vybrať model, pre ktorý platí $C_p \leq p$ a z hľadiska odhadov parametrov vybrať model, spĺňajúci podmienku $C_p \leq 2p - p_{full} + 1$. Podľa tohto kritéria je na prvej priečke model s desiatimi vysvetľujúcimi premennými uvedenými v stĺpci *Variables in model*. Podľa korigovaného koeficienta determinácie sa za najlepší model považuje model s najväčšou hodnotou tejto obľúbenej charakteristiky. Ostatné štatistiky by mali so zväčšujúcou kvalitou regresného modelu znižovať svoje hodnoty. Model, ktorý sme vybrali predchádzajúcimi metódami, sa podľa Mallowsvej štatistiky ($C_p = 5,305$) nachádza na 8. mieste v rebríčku kvality. Korigovaný koeficient determinácie má pre tento model len o jednu tisícinu (0,1 %) menšiu hodnotu ako pre model (11 vysvetľujúcich premenných) s najväčšou hodnotou tejto štatistiky. Podľa Swarzovho Bayesovského kritéria (SBC) je dokonca tento model najlepší. A aj ostatné charakteristiky poukazujú na dobrú kvalitu tohto modelu a na malé rozdiely v kvalite v porovnaní s ostatnými modelmi. Po zohľadnení výsledkov všetkých troch metód výberu najlepšej podmnožiny vysvetľujúcich premenných sme vybrali model s ôsmimi nezávislými premennými: *Nájomné, Počet izieb, Veľkosť domácnosti, Obydlie1, Vlastníctvo, Obydlie3, Počítač, Telefón*.

4. Diagnostikovanie vplyvných pozorovaní

Vplyvné pozorovania v regresnej analýze vznikajú najmä v dôsledku netypických kombinácií vysvetľujúcich premenných. Keďže väčšinou ide o viacrozmerné pozorovania, ich diagnostikovanie na základe grafickej analýzy je obtiažne. Na ich odhalenie sa preto používajú techniky, ktoré posudzujú dôsledok vypustenia i -teho pozorovania na regresné charakteristiky (odhady parametrov, vyrovnané hodnoty, odhady kovariančnej matice, atď.). Dôležitým nástrojom hodnotenia vplyvu jednotlivých pozorovaní na regresné charakteristiky je projekčná matica, resp. jej diagonálne prvky, ktoré sa nazývajú projekčné h -prvky. Tie sa označujú aj ako vplyv, efekt (leverage). Pozorovania s vysokou hodnotou projekčného h -prvku (odporúča sa $h_{ii} > \frac{3p}{n}$), sa považujú za extrémne (to ešte neznamená vplyvné).

Z projekčných h -prvkov vychádzajú skutočné charakteristiky pre identifikovanie vplyvných pozorovaní. My sme použili Cookovu štatistiku D_i , Jackknife rezíduá alebo navonok studentizované rezíduá (e_{ji} - *RSTUDENT*) a $DFFIT_i$. Čím je absolútna hodnota týchto charakteristík väčšia, tým je pozorovanie vplyvnejšie a väčšou mierou vplyva na odhady a induktívne úsudky používané v regresnej a korelačnej analýze. Preto sme analyzovaný súbor usporiadali podľa veľkosti Cookovej štatistiky (výstup 5). Ako značne vplyvné sme diagnostikovali pozorovanie s výrazne najväčšou hodnotou Cookovej štatistiky ($D_i = 2,405$), pre ktoré aj štatistika $DFFIT_i$ mala najväčšiu hodnotu ($DFFIT_i = 4,747$) a pozorovanie s extrémne vysokou hodnotou Jackknife rezídua ($e_{ji} = 8,000$). Tieto dve pozorovania sme z výberového súboru odstránili, a preto do ďalšej analýzy vstupovalo 523 pozorovaní závislej premennej – *Celkové mesačné náklady domácností na bývanie* a ôsmich vybraných vysvetľujúcich premenných.

Výstup 4: 10 najlepších modelov získaných metódou všetkých možných regresíí s kritériom najväčšieho upraveného koeficienta determinácie

Number in Model	C(p)	R-Square	Adjusted R-Square	Root MSE	AIC	BIC	Prediction Criterion	SBC	Variables in Model
10	4.4444	0.8688	0.8656	1306.88585	6138.6495	6141.5891	0.1382	6183.27408	X1 X14 X17 X10 X12 X13 X4 X6 X8 X18
9	4.6513	0.8681	0.8652	1308.83086	6138.9447	6141.6877	0.1382	6179.51256	X1 X14 X17 X10 X12 X13 X6 X8 X18
10	4.7139	0.8687	0.8655	1307.31574	6138.9303	6141.8550	0.1382	6183.55495	X1 X14 X17 X10 X12 X13 X4 X6 X8 X19
9	5.0175	0.8680	0.8651	1309.41288	6139.3244	6142.0490	0.1384	6179.89223	X1 X14 X17 X10 X12 X13 X6 X8 X19
9	5.0744	0.8679	0.8651	1309.50330	6139.3834	6142.1052	0.1384	6179.95121	X1 X14 X17 X10 X12 X13 X4 X6 X8
11	5.1563	0.8692	0.8657	1306.39952	6139.3039	6142.4077	0.1384	6187.98533	X1 X14 X17 X10 X12 X13 X15 X4 X6 X8 X18
10	5.2166	0.8685	0.8654	1308.11759	6139.4540	6142.3507	0.1384	6184.07860	X1 X14 X17 X10 X12 X13 X15 X6 X8 X18
8	5.3052	0.8672	0.8647	1311.47160	6139.6888	6142.2385	0.1385	6176.19985	X1 X14 X17 X10 X12 X13 X6 X8
9	5.7688	0.8677	0.8649	1310.60626	6140.1024	6142.7895	0.1386	6180.67020	X1 X14 X17 X10 X12 X13 X4 X6 X19
9	5.7884	0.8677	0.8649	1310.63739	6140.1226	6142.8088	0.1386	6180.69048	X1 X14 X17 X10 X12 X13 X4 X6 X18

Výstup 5: Diagnostika vplyvných pozorovaní

	• cookd_N na bývanie	• rstudent_N na bývanie	• dffits_N na bývanie	• h_N na bývanie	• student_N na bývanie
517	0.027507289	-1.946414254	-0.498902272	0.0616489594	-1.941176142
518	0.0327801995	8.0001353021	0.5753634815	0.0051457583	7.5523496653
519	0.0494494453	-2.492070779	-0.670476653	0.0674987885	-2.479583171
520	0.0561119758	2.2717498723	0.7134984178	0.0897859795	2.2626455889
521	0.0587905927	-2.332341571	-0.730525838	0.0893394695	-2.322371783
522	0.0681106736	2.9364280013	0.7887024446	0.0672876538	2.9149761938
523	0.1337406378	-1.885381114	-1.099829515	0.2538942314	-1.880731211
524	0.17461382	2.1513332808	1.2580041713	0.2548097032	2.1438094253
525	2.4054649372	4.6971884252	4.7468856533	0.5052621201	4.604155489

5. Posúdenie vplyvu vysvetľujúcich premenných na náklady domácností na bývanie

Lineárny regresný model vystihujúci závislosť celkových mesačných nákladov domácností na bývanie od vysvetľujúcich premenných: *Nájomné*, *Počet izieb*, *Veľkosť domácnosti*, *Obydlie1*, *Obydlie3*, *Vlastníctvo*, *Počítač*, *Telefón* môžeme, podľa tabuľky analýzy rozptylu závislej premennej vo výstupe 6, považovať na ľubovoľnej bežne používanej hladine významnosti za štatisticky významný ($p < 0,0001$). Vieme ním vysvetliť 76,32 % variability nákladov na bývanie. Zvyšných 23,68 % variability nákladov na bývanie spôsobujú faktory neuvažované v regresnom modeli a náhodné vplyvy.

Výstup 6: Analýza rozptylu pre regresný model

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	8	2 451 856 017	306 482 002	207.10	<.0001
Error	514	760 644 070	1 479 852		
Corrected Total	522	3 212 500 088			

Root MSE	1216.49179	R-Square	0.7632
Dependent Mean	5266.24192	Adj R-Sq	0.7595
Coeff Var	23.09981		

Odhady parametrov tohto modelu získané metódou najmenších štvorcov sú vo výstupe 7 v stĺpci *Parameter Estimate*. Všetky regresné koeficienty modelu sú po odstránení dvoch vplyvných pozorovaní štatisticky významné, a to nielen na požadovanej úrovni významnosti 0,1, ale aj na úrovni 0,05 (všetky p -hodnoty v stĺpci $Pr > |t|$ sú menšie ako 0,05). To znamená, že prínos všetkých vysvetľujúcich premenných k vysvetleniu variability nákladov domácností na bývanie je podstatný.

Predstavu o veľkosti prínosu poskytujú stĺpce *Squared Partial Corr Type I*, *Squared Semi-partial Corr Type I*, *Squared Partial Corr Type II*, *Squared Semi-partial Corr Type II*. V stĺpci *Squared Partial Corr Type I* je koeficient determinácie medzi vysvetľovanou premennou a zodpovedajúcou vysvetľujúcou premennou pri eliminácii vplyvu premenných, ktoré sú v predchádzajúcich riadkoch. To znamená, že premenná *Nájomné* vysvetľuje 60,806 % variability *Nákladov domácností na bývanie*. Zo zvyšnej (nevysvetlenej) variability vysvetľuje premenná *Počet izieb* 26,85 %, čo je prínos k vysvetleniu variability premennej *Náklady domácností na bývanie* 10,524 % (uvedený v stĺpci *Semi-partial Corr Type I*). Tretia vysvetľujúca premenná *Veľkosť domácnosti* vysvetľuje z tej časti variability premennej *Náklady domácností na bývanie*, ktorú nevysvetlili premenné *Nájomné* a *Počet izieb* 6,406 %, čo predstavuje dodatočný prínos k vysvetleniu variability závislej premennej 1,837 %. Hodnoty v ďalších dvoch stĺpcoch (označené *Type II*) môžeme interpretovať obdobne, s tým rozdielom, že eliminujeme vplyv premenných, ktoré sú uvedené v ostatných riadkoch tabuľky (nie v predchádzajúcich riadkoch ako pri *Type I*).

Významnosť prínosu jednotlivých premenných potvrdzujú aj 90 %-né intervaly spoľahlivosti regresných koeficientov. Podľa nich môžeme s 90 %-nou pravdepodobnosťou napríklad tvrdiť, že so zvýšením súčasného nájomného o 100 Sk sa mesačné náklady na bývanie zvýšia v priemere o viac ako 81,4 Sk a súčasne o menej ako 89,5 Sk, za predpokladu, že hodnoty ostatných premenných sa nezmenia. Keďže z množiny premenných, ktoré charakterizujú typ

Výstup 7: Posúdenie vplyvu vysvetľujúcich premenných na náklady domácností na bývanie regresnými a korelačnými charakteristikami

Parameter Estimates												
Variable	Parameter Estimate	Standard Error	t Value	Pr > t	Squared Semi-partial Corr Type I	Squared Partial Corr Type I	Squared Semi-partial Corr Type II	Squared Partial Corr Type II	Tolerance	Variance Inflation	90% Confidence Limits	
Intercept	-1198.706	461.719	-2.60	0.0097	.	.	.	.	.	0	-1959.540	-437.875
Nájomné	0.855	0.025	34.54	<.0001	0.60806	0.60806	0.54964	0.69892	0.94730	1.05563	0.814	0.895
Počet izieb	597.960	74.620	8.01	<.0001	0.10524	0.26850	0.02958	0.11106	0.67103	1.49025	475.000	720.921
Veľkosť domácnosti	674.999	94.853	7.12	<.0001	0.01837	0.06406	0.02333	0.08969	0.72967	1.37049	518.698	831.300
Obydlie1	-1748.860	310.895	-5.63	<.0001	0.01803	0.06721	0.01458	0.05799	0.93088	1.07425	-2261.159	-1236.560
Obydlie3	-1770.952	611.144	-2.90	0.0039	0.00360	0.01439	0.00387	0.01607	0.99817	1.00183	-2778.010	-763.895
Vlastníctvo	1381.405	371.486	3.72	0.0002	0.00568	0.02302	0.00637	0.02620	0.91460	1.09337	769.261	1993.549
Telefón	600.585	243.071	2.47	0.0138	0.00213	0.00883	0.00281	0.01174	0.94513	1.05806	200.048	1001.123
Počítač	-242.925	113.240	-2.15	0.0324	0.00212	0.00887	0.00212	0.00887	0.88757	1.12667	-429.524	-56.326

Výstup 8 : Diagnostika multikolinearity

Number	Proportion of Variation										
	Eigenvalue	Condition Index	Intercept	Nájomné	Počet izieb	Veľkosť domácnosti	Obydlie1	Obydlie3	Vlastníctvo	Telefón	Počítač
1	5.93566	1.00000	0.00036178	0.00816	0.00180	0.00187	0.00104	0.00030596	0.00056472	0.00135	0.00792
2	1.01134	2.42262	2.671965E-7	0.00857	0.00006489	0.00003483	0.46115	0.44276	0.00001597	0.00001778	0.00460
3	0.97441	2.46811	0.00000684	0.00395	0.00023238	0.00009555	0.39181	0.55481	0.00003808	0.00005047	0.00712
4	0.45173	3.62490	0.00020400	0.30347	0.00015712	0.00003439	0.08390	0.00001692	0.00032533	0.00027666	0.69660
5	0.43705	3.68528	0.00144	0.66457	0.00327	0.00576	0.01725	0.00056244	0.00236	0.00376	0.23089
6	0.09241	8.01432	0.00992	0.00004082	0.17403	0.27601	0.00102	0.00020406	0.01594	0.14288	0.01618
7	0.04909	10.99553	0.00272	0.00763	0.77740	0.61214	0.00937	0.00093056	0.00086067	0.00037235	0.00154
8	0.03959	12.24460	0.02449	0.00137	0.02377	0.02362	0.00040102	0.00004984	0.20551	0.65162	0.02684
9	0.00872	26.09462	0.96085	0.00225	0.01927	0.08043	0.03405	0.00035554	0.77439	0.19968	0.00832

obydli sa do modelu nezaradila premenná *Obydlie 2*, náklady na bývanie v byte, ktorý je v bytovom dome s menej ako 10 bytmi a náklady na bývanie v byte, ktorý je v bytovom dome s najmenej 10 bytmi nemôžeme považovať za rôzne. Na druhej strane, s 90 %-nou spoľahlivosťou má rodina, ktorá obýva samostatný dom v porovnaní s rodinou, ktorá obýva bytový dom, mesačné náklady na bývanie v priemere nižšie najmenej o 1 236,56 Sk a najviac o 2 261,16 Sk, za predpokladu, že sa eliminuje vplyv ostatných vysvetľujúcich premenných. Rodina, ktorá využíva iný druh bývania má tieto náklady v porovnaní s rodinou, ktorá býva v bytovom dome s 90 %-nou pravdepodobnosťou nižšie v priemere o 763,90 Sk až o 2 778,01 Sk, za predpokladu, že porovnávame rodiny s rovnakými hodnotami ostatných vysvetľujúcich premenných.

6. Posúdenie stupňa multikolinearity

Pod pojmom multikolinearita (z angl. multicollinearity), resp. kolinearita (z angl. collinearity) rozumieme závislosť medzi vysvetľujúcimi premennými. V klasickom lineárnom regresnom modeli nemôže byť žiadna vysvetľujúca premenná lineárnou kombináciou iných vysvetľujúcich premenných. Ak tento predpoklad nie je splnený, odhad parametrov modelu metódou najmenších štvorcov nemá jednoznačné riešenie. Vysvetľujúce premenné sú zväčša navzájom aspoň čiastočne lineárne závislé. So zväčšujúcim sa stupňom multikolinearity sa zväčšujú rozptyly odhadov parametrov modelu a komplikuje sa alebo dokonca sa znemožňuje rozumná interpretácia výsledkov regresnej a korelačnej analýzy. Z hľadiska výpovednej schopnosti regresného modelu preto musíme posúdiť, či stupeň multikolinearity je alebo nie je nežiadúco vysoký.

Na posúdenie toho, či v modeli je nežiadúco vysoká multikolinearita a ak je, ktorá(é) premenná(é) ju spôsobuje(ú), sme použili štatistiky *Tolerance*, VIF_j (*Variance Inflation factor*) vo výstupe 7 a vlastné čísla (*Eigenvalue*), index podmienenosti (*Condition Index*), štatistiky $p_{jj'}$ (*Proportion of Variation*) vo výstupe 8. *Tolerance* je výberový koeficient neurčitosti medzi príslušnou vysvetľujúcou premennou na jednej strane a ostatnými vysvetľujúcimi premennými na strane druhej. Ideálna, no prakticky nereálna, je hodnota $Tolerance=1$ (lineárna nezávislosť j -tej vysvetľujúcej premennej od ostatných vysvetľujúcich premenných). Čím je hodnota tejto charakteristiky bližšia k nule, tým zodpovedajúca vysvetľujúca premenná väčšou mierou prispieva k multikolinearite. Obrátená hodnota štatistiky *Tolerance* je VIF_j , ktorej hodnoty väčšie ako 10 signalizujú prítomnosť silnej multikolinearity. Na základe oboch mier stupeň multikolinearity v analyzovanom regresnom modeli nie je nežiadúco vysoký. Stupeň závislosti medzi vysvetľujúcimi premennými je príliš vysoký aj v prípade, že niektoré vlastné číslo λ_i je blízke nule a index podmienenosti η_j je vysoký (väčší ako 30). Ani podľa indexu podmienenosti nie je multikolinearita problematická. V prípade, že by sa preukázala nežiadúco vysoká multikolinearita, potom v riadku s najväčším indexom podmienenosti (najmenším vlastným číslom) hodnoty *Proportion of Variation* väčšie ako 0,5 odhaľujú množinu vysvetľujúcich premenných, ktoré najväčšou mierou prispievajú k multikolinearite.

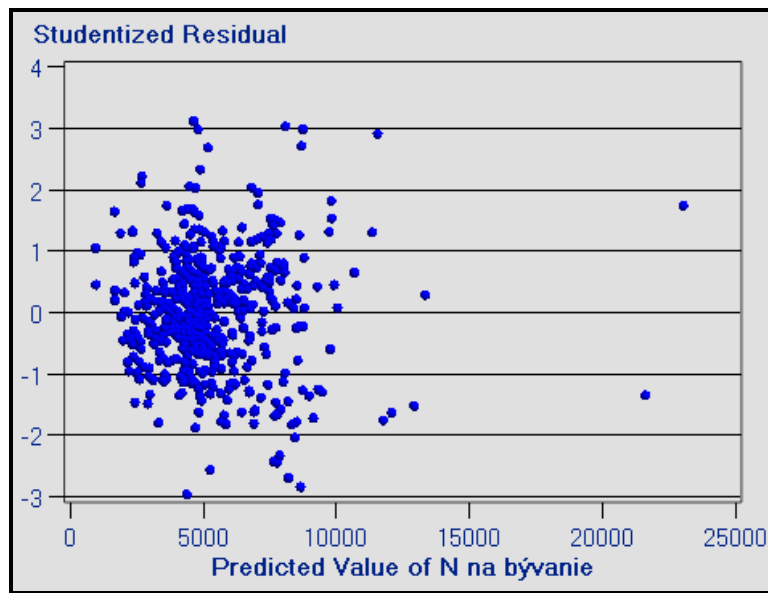
7. Overenie predpokladov klasického lineárneho regresného modelu

Klasický lineárny regresný model kladie na náhodnú zložku tieto podmienky:

- Stredná hodnota náhodných chýb ε_i sa rovná nule: $E(\varepsilon_i) = 0$, pre všetky $i = 1, 2, \dots, n$.
- Rozptýlenosť (variabilita) hodnôt závislej premennej Y pre ktorúkoľvek úroveň vektora $\mathbf{x}_0$ je rovnaká, t. j. rozptyl náhodnej zložky je konštantný (homoskedastický): $D(\varepsilon_i) = \sigma_\varepsilon^2$.

- c) Rozdelenie pravdepodobnosti náhodných chýb je normálne so strednou hodnotou 0 a rozptylom $\sigma_\varepsilon^2: \varepsilon_i \sim N(0; \sigma_\varepsilon^2)$.
- d) Náhodné chyby sú pre ľubovoľnú dvojicu hodnôt vektora vysvetľujúcich premenných $\mathbf{x}_i \neq \mathbf{x}_j$ vzájomne nezávislé: $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$ pre všetky $i \neq j$.

Prvý predpoklad sa neoveruje a na overenie ostatných predpokladov použijeme grafickú analýzu rezíduí a štatistické testy. Pri grafickej analýze vychádzame z grafu vyrovnaných hodnôt závislej premennej verzus studentizované rezíduá na obr. 1.



Obr. 1: Vyrovnané hodnoty závislej premennej verzus studentizované rezíduá

Variabilita rezíduí sa so zmenami vyrovnaných hodnôt výrazne nemení, preto nepredpokladáme porušenie homoskedasticity náhodnej zložky. Studentizované rezíduá sa koncentrujú okolo osi y a ich frekvencia sa v smeroch od nulovej úrovne znižuje. V intervale $(-2; 2)$, v ktorom je zastúpených 95,45 % hodnôt premennej s normovaným normálnym rozdelením, sa nachádza prevažná časť studentizovaných rezíduí a v intervale $(-3; 3)$, v ktorom je 99,73 % hodnôt premennej s normovaným normálnym rozdelením, sa s výnimkou dvoch až štyroch nachádzajú všetky studentizované rezíduá. Tieto zistenia podporujú predpoklad o normálnom rozdelení náhodnej zložky. Keďže rezíduá v grafickom poli nevykazujú systematický priebeh, nič nenaznačuje porušenie predpokladu o nezávislosti náhodných chýb.

Na overenie homoskedasticity náhodnej zložky sme ešte použili testy homogenity (*Leveneov*, *Brown-Forsytheov* a *Bartlettov*), ktoré sú uvedené vo výstupe 9. Vo všetkých sa testuje nulová hypotéza

$$H_0 : \text{náhodná zložka je homoskedastická}$$

oproti alternatívnej hypotéze

$$H_1 : \text{náhodná zložka je heteroskedastická}.$$

Na základe p -hodnôt ($p = 0,1991$; $p = 0,0823$; $p = 0,1202$), na hladine významnosti 0,05 ($p > \alpha$) nemôžeme zamietnuť nulovú hypotézu o konštantnom rozptyle náhodnej zložky.

Výstup 9: Overenie homoskedasticity náhodnej zložky

Levene's Test for Homogeneity of residual_N na bývanie2 Variance ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
F	1	7.285E12	7.285E12	1.65	0.1991
Error	520	2.292E15	4.407E12		

Brown and Forsythe's Test for Homogeneity of residual_N na bývanie2 Variance ANOVA of Absolute Deviations from Group Medians					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
F	1	1553156	1553156	3.03	0.0823
Error	520	2.6655E8	512590		

Bartlett's Test for Homogeneity of residual_N na bývanie2 Variance			
Source	DF	Chi-Square	Pr > ChiSq
F	1	2.4142	0.1202

Na overenie tretieho predpokladu, sme na úrovni významnosti 0,05 testovali nulovú hypotézu

$H_0 : \varepsilon_i \sim N(0; \sigma_\varepsilon^2)$, t. j. náhodná zložka má normálne rozdelenie s parametrami 0 a σ_ε^2 oproti alternatívnej hypotéze

$H_1 : \text{non } H_0$, t. j. náhodná zložka nemá takéto rozdelenie.

Na základe *Kolmogorov-Smirnovho testu*, *Cramer-von-Misesovho testu* aj *Anderson-Darlingovho testu* vo výstupe 10, nezamietame nulovú hypotézu ($p > \alpha = 0,05$), a teda môžeme predpokladať normálne rozdelenie náhodnej zložky.

Výstup 10: Overenie normálneho rozdelenia náhodnej zložky

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.04853736	Pr > D	0.168
Cramer-von Mises	W-Sq	0.27008779	Pr > W-Sq	0.169
Anderson-Darling	A-Sq	1.78313694	Pr > A-Sq	0.123

Nakoniec sme ešte overili nezávislosť náhodných chýb. Využili sme ponuku z lineárnej regresie v SAS EG, ktorá poskytuje *Durbin-Watsonovu štatistiku* a *koeficient autokorelácie 1. stupňa* uvedené vo výstupe 11.

Výstup 11: Overenie nezávislosti náhodných chýb

Durbin-Watson D	1.902
Number of Observations	524
1st Order Autocorrelation	0.046

Durbin-Watsonova štatistika $D - W = 1,902$ sa približuje k ideálnej hodnote 2 a je v oblasti

prijatia nulovej hypotézy ($d_U; 4 - d_U$), kde d_U má pre $n = 523$, $k = 8$ a $\alpha = 0,05$ hodnotu $d_U = 1,884$. Preto na hladine významnosti 0,05 môžeme predpokladať nezávislosť náhodných chýb. Výsledok tohto testu potvrdzuje aj hodnota koeficienta autokorelácie $r_1 = 0,046$, ktorá hovorí o veľmi slabej priamej lineárnej závislosti medzi radom rezíduí e_i a radom rezíduí posunutých o jedno pozorovanie e_{i-1} . Rozdiel medzi hodnotou koeficienta autokorelácie $r_1 = 0,046$ vo výbere a nulovou hodnotou ρ_1 - koeficienta autokorelácie v základnom súbore je však štatisticky nevýznamný.

8. Záver

V príspevku sme modelovali závislosť mesačných nákladov domácností na bývanie od vybraných faktorov. Metódami najlepšej podmnožiny vysvetľujúcich premenných sme z pôvodnej množiny kandidujúcich vysvetľujúcich premenných vybrali tie, ktoré relevantne vplyvajú na náklady domácností na bývanie. Po odstránení vplyvných pozorovaní sme metódou najmenších štvorcov odhadli klasický lineárny regresný model. Metódami regresnej a korelačnej analýzy sme ohodnotili vplyv vysvetľujúcich premenných na náklady domácností na bývanie. Keďže sa nepotvrdila nežiadúco vysoká multikolinearita a ani porušenie predpokladov o náhodnej zložke, odhadnutý regresný model je vhodný na posúdenie vzťahu medzi analyzovanými premennými, ako aj na prognostické účely.

Príspevok je súčasťou riešenia projektu VEGA č. 1/4586/07: Modelovanie sociálnej situácie obyvateľstva a domácností v SR a jej regionálne a medzinárodné porovnania.

9. Literatúra

- BAKYTOVÁ, H. – BODJANOVÁ, S. – RUBLÍKOVÁ, E. 1990. Viacrozmerná analýza. Bratislava: ES VŠE, 1990. 289 s. ISBN 80-225-0184-0.
- DRAPER, N. – SMITH, H. 1985. Applied Regression Analysis. New York: J. Wiley & Sons, 1985. 709 s. ISBN 0-471-02995-5.
- HEBÁK, P. – HUSTOPECKÝ, J. – MALÁ, I. 2005. Vícerozměrné statistické metody [2]. Praha: Informatorium, 2005. 239 s. ISBN 80-7333-036-9.
- KVANLI, A. H. – GUYNES, C. S. – PAVUR, R. J. 1992 Introduction to Business Statistics: a Computer Integrated Approach. (Third Edition), St. Paul, New York, Los Angeles, San Francisco: West publishing company, 1992. 930 s. ISBN 0-314-89946-4.
- MILES, J. – SHEVLIN, M. 2001. Applying regression and correlation. London: SAGE publications, 2001. 253 s. ISBN 0-7619-6229-8.
- NETER, J. – WASSERMAN, W. 1974. Applied linear statistical models. Homewood: IRWIN, 1974. 842 s. ISBN 0-256-01498-1.
- PFAFFENBERGER, R. C. – PATTERSON, J. H. 1987. Statistical Methods for Business and Economics. (Third Edition), Homewood: IRWIN, 1987. 1 246 s. ISBN 0-256-03664-0.

Adresa autora:

Erik Šoltés, Mgr., PhD.
Katedra štatistiky, FHI EU,
Dolnozemska cesta 1
852 35 Bratislava
e-mail: soltes@euba.sk

Logistická regresia a jej využitie v ekonomickej praxi

Iveta Stankovičová¹

Abstract: Binomial (binary) logistic regression is a form of regression, which is used when the dependent is a dichotomy and the independents are of any type. Multinomial logistic regression exists to handle the case of dependents with more classes than two. Logistic regression can be used to predict a dependent variable on the basis of continuous and/or categorical independents and to determine the percent of variance in the dependent variable explained by the independents; to rank the relative importance of independents; to assess interaction effects; and to understand the impact of covariate control variables. Logistic regression applies maximum likelihood estimation after transforming the dependent into a logit variable (the natural log of the odds of the dependent occurring or not). In this way, logistic regression estimates the probability of a certain event occurring.

Many procedures in SAS/STAT[®] can be used to perform logistic regression analysis: CATMOD, GENMOD, LOGISTIC, and PROBIT. Each procedure has special features that make it useful for certain applications. For most applications, PROC LOGISTIC is the preferred choice.

Key words: logistic regression, logit, odds ratio, maximum likelihood estimation, SAS system.

1. Úvod

Logistická regresia sa v značnej miere uplatňuje v lekárskej a farmaceutickej praxi (napr. skúmanie pravdepodobnosti ochorenia pacienta pri známych hodnotách vybraných laboratórnych testov). V marketingu (hlavne pri prieskume trhu) je táto metóda veľmi užitočná, pretože sa často snažíme predvídať diskretnú akciu (napr. odozvu na ponuku: 1=odpovedal, 0=neodpovedal; splácanie pôžičky: 1=splatil včas, 2=splatil s oneskorením, 3=nesplatil). Nemalé využitie má však aj vo finančnej praxi. Táto analytická metóda má veľmi dobré uplatnenie pri rozhodovaní o poskytnutí úveru, odporúčení niektorej z bankových služieb, či type poistenia, používa sa tiež na detekciu podvodov a podobne.

Logistická regresia je metóda vhodná na modelovanie jednostrannej závislosti medzi premennými. Ako závislá premenná v modeli vystupuje kategóriálna premenná a ako vysvetľujúce môžu byť premenné spojitého alebo kategóriálneho typu. Ide o modelovanie podmienenej pravdepodobnosti jednej obmeny závislej kategóriálnej premennej v závislosti od iných premenných. V príspevku sa obmedzíme na výklad logistických modelov s binárnou závislou premennou, pretože problematika je veľmi rozsiahla.

2. Model logistickej regresie

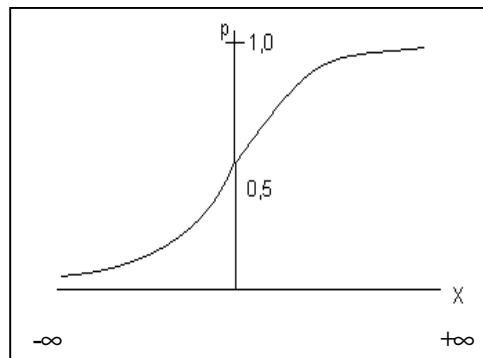
Logistická regresia je veľmi podobná lineárnej regresii. Aj logistická regresia je založená na štatistickom rozdelení a patrí medzi robustné nástroje pre tvorbu modelov. Hlavný rozdiel je v tom, že závislá premenná nie je spojitá, ale je diskretná, resp. kategóriálna. Aby sa dala použiť regresia, tak sa závislá premenná transformuje na spojitú

¹ Ing. Iveta Stankovičová, PhD., Katedra informačných systémov, FM UK v Bratislave

hodnotu, ktorá je funkciou pravdepodobnosti výskytu udalosti. Logistická regresia je špeciálnym prípadom zovšeobecného lineárneho modelu (GLM = General Linear Model).

Označme ako $\mathbf{X}$ vektor nezávislých premenných. Označme ako Y binárnu závislú premennú, ktorá nadobúda dve možné hodnoty 0 (neúspech, resp. želaná udalosť nenastala) a 1 (úspech, resp. želaná udalosť nastala). Označme symbolom p pravdepodobnosť výskytu želanej udalosti ($Y=1$), teda hodnotu, ktorú modelujeme. Potom $p=P(Y=1/X)$ je podmienená pravdepodobnosť výskytu želanej udalosti za podmienky výskytu vektora nezávislých premenných $\mathbf{X}$. Podiel $p/(1-p)$ vyjadruje šance (odds) výskytu želanej udalosti (úspechu) k výskytu neželanej udalosti (neúspechu). Šance a pravdepodobnosť predstavujú tie isté informácie, len v iných podobách, resp. inými slovami.

Vzťah medzi pravdepodobnosťou p a vysvetľujúcou premennou $\mathbf{X}$ nie je lineárny. Grafickým zobrazením závislosti je logistická sigmoidná krivka, tzv. S-krivka (Obrázok 1).



Obrázok 1: S-krivka

Pre finálnu transformáciu závislej premennej na spojitú premennú sa vypočíta logaritmus šancí, tzv. $\text{logit}(p)=\ln(p/(1-p))$, resp. $\text{logit}(p)=\log(p/(1-p))$.² Vzťah medzi logitom a vektorom vysvetľujúcich premenných má už lineárny charakter. Rovnica logistického modelu má nasledovný tvar:

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k. \quad (1)$$

Vzťah medzi pravdepodobnosťou a vektorom vysvetľujúcich premenných dostaneme spätnou transformáciou a má nelineárny charakter. Je to druh exponenciálnej funkcie:

$$p = \frac{e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}}{1 + e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}}. \quad (2)$$

Logitovou transformáciou (Obrázok 2) dostávame z nelineárnej závislosti závislosť lineárnu. Kým hodnoty pravdepodobností p_i sú z intervalu $(0,1)$, hodnoty podielov $p_i/(1-p_i)$ sú nezáporné hodnoty, tak hodnoty logitov môžu nadobúdať akékoľvek reálne hodnoty z intervalu $(-\infty, \infty)$.

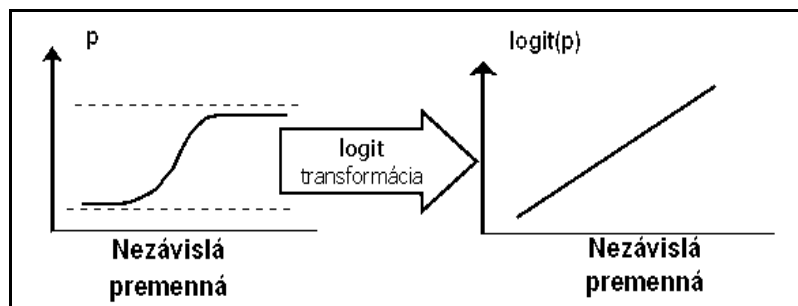
Modely logistickej regresie môžu obsahovať rôzne druhy vysvetľujúcich premenných (prediktorov X_i). Vo všeobecnosti sú to:

- kvantitatívne premenné (napr. VEK, PRIJEM),
- polynómy kvantitatívnych premenných (napr. VEK, VEK2, VEK3),
- transformácie kvantitatívnych premenných (napr. $\log(\text{PRIJEM})$),

² Pri logitovej transformácii je možné použiť prirodzený logaritmus ($\ln$) alebo aj dekadický ($\log$) logaritmus. Výsledky sú analogické, preto označenie „log“ znamená všeobecný symbol pre logaritmus, tzn. či dekadický alebo prirodzený. V praxi sa bežnejšie používa prirodzený logaritmus.

- kvalitatívne premenné (napr. POHLAVIE), ktoré vystupujú väčšinou v úlohe formálnych (dummy) premenných,
- interakčné premenné (napr. VEK*PRIJEM, POHLAVIE*PRIJEM).

Charakter vysvetľujúcich premenných je podstatný pre konštrukciu modelu, pre odhad jeho parametrov, hodnotenie jeho kvality a aj využitie. Dôležité je, či sú kombinácie vysvetľujúcich premenných jedinečné (tak ako je to pri spojitých premenných), a tak môžeme na konštrukciu logistického modelu použiť netriedené údaje. Keď uvažujeme len s vysvetľujúcimi premennými kategoriálneho typu, tak je možné vytriediť údaje do viacrozmernej kontingenčnej tabuľky, a potom ich použiť pre modelovanie. Logistická regresia sa v prípade, že nezávislou premennou je len jedna kategoriálna premenná, redukuje na analýzu kontingenčnej tabuľky.



Obrázok 2: Logitová transformácia

3. Metódy odhadu parametrov modelu logistickej regresie

Na odhad parametrov modelu logistickej regresie sa dnes najčastejšie používa metóda maximálnej vierohodnosti (ML)³. Táto metóda sa mohla začať používať až s rozvojom výpočtovej techniky, lebo je výpočtovo veľmi náročná. Je vhodná na odhad parametrov nelineárnych modelov, kde patrí aj logistická regresia. V minulosti sa na odhad parametrov logistických modelov používala aj metóda diskriminačnej funkčnej analýzy, ktorá vyžaduje podmienku normality rozdelenia pre vysvetľujúce premenné (tak ako metóda najmenších štvorcov). V prípade, že niektorá vysvetľujúca premenná je kategoriálna, táto metóda dáva skreslené (resp. vychýlené) výsledky, zvyčajne sa to prejaví tak, že odhady pomerov šancí (odds ratio) sú veľmi vysoké. Metóda maximálnej vierohodnosti (ML) nemá žiadne požiadavky na nezávislé premenné a tie môžu byť nominálne, poradové (ordinálne), alebo intervalové (diskrétné alebo spojité).

Poznáme dva typy metód maximálnej vierohodnosti (ML): podmienenú (conditional) a nepodmienenú (unconditional) ML. Tieto dva typy vyžadujú aj rôzne počítačové programy. Systém SAS a tiež jeho interaktívna aplikácia SAS Enterprise Guide (ďalej len SAS EG) v procedúre LOGISTIC používa nepodmienenú metódu⁴.

Nepodmienená metóda je štatistikmi preferovaná vtedy, keď počet parametrov modelu je relatívne nízky k rozsahu vzorky dát (napr. 12 parametrov k rozsahu vzorky $n = 700$). Podmienená metóda je preferovaná v opačnom prípade, kedy počet parametrov logistického modelu je vysoký k rozsahu vzorky (napr. 103 parametrov a len $n = 200$), čo sa môže vyskytnúť v modeloch s interakčnými premennými.

³ ML je skratka anglického výrazu „maximum likelihood method“.

⁴ Nepodmienená metóda maximálnej vierohodnosti sa používa aj v softvéroch SPSS, S+ a BMDP.

Odhad parametrov logistického modelu získame z výberovej vzorky dát metódou maximálnej vierohodnosti tak, že maximalizujeme vierohodnostnú funkciu L (resp. jej upravený logaritmus $-2\log L$) výberových údajov vzhľadom na neznáme parametre modelu. Odhad funkcie L získame zo vzorky podľa nasledovného vzorca:

$$-2\log L = -2 \sum_{i=1}^n \log(\hat{p}_i) , \quad (3)$$

kde $\hat{p}_i$ sú bodové odhady podmienených pravdepodobností modelovanej hodnoty pre všetky kombinácie vysvetľovaných premenných získané na základe výberových dát.

Vierohodnostná funkcia L je vlastne funkciou neznámych parametrov v logistickom modeli a môžeme ju tiež zapísať v nasledovnom tvare:

$$L = L(\theta)$$

$$\theta = (\theta_1, \theta_2, \theta_3, \dots, \theta_p) = (\beta_0, \beta_1, \beta_2, \dots, \beta_k)$$

kde p znamená počet parametrov a k počet vysvetľujúcich premenných v logistickom modeli. Vierohodnostná funkcia L predstavuje vlastne zjednotenie pravdepodobností získaných z pozorovaných dát pre modelovanú hodnotu závislej premennej.

Metóda maximálnej vierohodnosti (ML) hľadá maximum funkcie L . Pri nepodmienenej ML ide vlastne o parciálnu deriváciu nasledovnej funkcie:

$$L_N = \prod_{l=1}^m P(X_l) \prod_{l=m+1}^n [1 - P(X_l)] , \text{ pre } l - \text{té pozorovanie}$$

$$P(X) = \hat{p} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}}$$

$$\max[\log(L(\theta))] \Rightarrow \frac{\partial \log(L(\theta))}{\partial \theta_j} = 0, j = 1, 2, \dots, p$$

kde m znamená počet výskytov modelovanej hodnoty závislej premennej vo vzorke dát n .

4. Testovanie logistického modelu

Po odhade parametrov logistického modelu metódou maximálnej vierohodnosti zvyčajne nasleduje testovanie významnosti (resp. vhodnosti) logistického modelu ako celku a tiež testovanie významnosti jeho jednotlivých parametrov. Pri logistickej regresii sa k tomuto účelu využíva odhadnutá hodnota L štatistiky a štandardné chyby odhadov jednotlivých parametrov. Na L štatistike je založený test vierohodnostným pomerom (likelihood ratio test) a z štandardných chýb vychádza Waldov test (Wald test). Obe testovacie procedúry dávajú približne rovnaké výsledky pre veľké vzorky dát, ale pre malé alebo stredné vzorky sa výsledky môžu odlišovať. Vo všeobecnosti sa dá povedať, že štatistickí preferujú pri malých a stredných vzorkách test vierohodnostným pomerom pred Waldovým testom.

Najskôr testujeme štatistickú významnosť modelu ako celku. Hypotézy sú formulované nasledovne:

H_0 : logistický model ako celok nie je významný, t.j. má parametre β_i rovné nule,

H_1 : logistický model ako celok je významný, t.j. nemá všetky parametre β_i rovné nule.

Požadujeme zamietnutie nulovej hypotézy. Na testovanie sa používa štatistika L (L =Likelihood), resp. jej transformácia $-2\log L$, ktorú získame na základe už vyššie uvedeného vzorca (3). Pretože odhadnuté pravdepodobnosti $\hat{p}_i$ sú čísla z intervalu $(0,1)$, tak po transformácii $-2\log L$ získame hodnoty z intervalu $(0,\infty)$. Požadujeme čo najväčšie L , ale opačne, $-2\log L$ čo najmenšie. Štatistika L predstavuje pre logistický model analógiu koeficientu determinácie r^2 (R-squared) pre klasický lineárny model.

Uvažujme, že pre vzorku dát sme odhadli parametre nasledovných štyroch logistických modelov a ich L štatistiky:

$$\begin{aligned}\text{logit } P_0(X) &= \beta_0 \\ \text{logit } P_1(X) &= \beta_0 + \beta_1 X_1 \\ \text{logit } P_2(X) &= \beta_0 + \beta_1 X_1 + \beta_2 X_2 \\ \text{logit } P_3(X) &= \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3\end{aligned}$$

Pre model s väčším počtom parametrov zvyčajne platí, že lepšie modeluje dáta a jeho L štatistika bude vyššia. Pre naše modely potom platí:

$$\begin{aligned}L_0 &\leq L_1 \leq L_2 \leq L_3 \\ \log L_0 &\leq \log L_1 \leq \log L_2 \leq \log L_3\end{aligned}$$

Pre transformáciu $-2\log L$ už však platí obrátený vzťah:

$$-2\log L_0 \geq -2\log L_1 \geq -2\log L_2 \geq -2\log L_3.$$

Pre testovanie hypotézy o významnosti odhadnutého logistického modelu ako celku sa používa test vierohodnostným pomerom (likelihood ratio test). Štatistika LR (likelihood ratio), ktorá sa slovensky môže nazývať aj pravdepodobnostný pomer, má chí-kvadrát rozdelenie s počtom stupňov voľnosti rovným rozdielu medzi počtom parametrov dvoch testovaných modelov.

Chceme porovnať prvé dva modely z nášho zoznamu: prvý len s absolútnym členom a druhý s jednou vysvetľujúcou premennou. Použijeme štatistiku $LR=L_0/L_1$, ktorá sa po logaritmickej transformácii rovná:

$$-2\log(LR) = -2\log\left(\frac{L_0}{L_1}\right) = -2\log L_0 - (-2\log L_1) \quad . \quad (4)$$

Model len s absolútnym členom má nižšiu hodnotu L štatistiky ako druhý model s jednou vysvetľujúcou premennou, preto hodnota transformovanej LR štatistiky je kladné číslo. V prípade, že obe hodnoty L štatistiky sú približne rovnaké, tak hodnota transformovanej LR štatistiky sa približne rovná nule. Vo všeobecnosti sa hodnoty transformovanej LR štatistiky môžu nachádzať v intervale $(0,\infty)$. Ak je hodnota transformovanej LR štatistiky blízka nule, testovaný model možno považovať za nevýznamný a v opačnom prípade, čím je hodnota transformovanej LR štatistiky vyššia, tým máme viac dôkazov o tom, že testovaný model lepšie vystihuje dáta ako model len s absolútnym členom.

Transformovaná LR štatistika sa používa tiež na testovanie tej skutočnosti, či pridanie ďalšej premennej do modelu zlepšuje vysvetlenie závislej premennej.

Iný spôsob ako otestovať významnosť logistického modelu je Waldov test. Waldova testovacia štatistika sa vypočíta nasledovne:

$$z_{\text{Wald}} = \frac{\hat{\beta}_j}{s_{\hat{\beta}_j}} . \quad (5)$$

Má približne normálne rozdelenie $N(0,1)$ pre veľkú vzorku. Štvorec Waldovej štatistiky (z_{Wald}^2) má však približne chí-kvadrát rozdelenie s jedným stupňom voľnosti. Pre veľké vzorky platí, že $LR \approx z_{\text{Wald}}^2$. Pre malé a stredné vzorky to nemusí platiť. Waldova štatistika sa používa hlavne pri porovnávaní modelov s počtom parametrov vyšším o jeden parameter.

5. Intervaly spoľahlivosti pre parametre logistického modelu

Poznáme dve metódy výpočtu intervalov spoľahlivosti pre parametre logistického modelu. Prvá je založená na tvare vierohodnostnej funkcie a druhá metóda predpokladá asymptoticky normálne rozdelenie pre odhady parametrov. Táto druhá metóda je výpočtovo menej náročná, lebo nevyžaduje iterácie, hoci pri malých vzorkách môže viesť ku skresleným výsledkom. Intervaly spoľahlivosti takto vypočítané sa volajú Waldove a nachádzajú sa aj vo výstupoch SAS EG.

Waldov 100(1- α)% interval spoľahlivosti pre parameter β_j vypočítame podľa vzorca:

$$\hat{\beta}_j \pm z_{1-\alpha/2} s_{\hat{\beta}_j} , \quad (6)$$

kde $\hat{\beta}_j$ sú maximálne vierohodné odhady parametrov β_j , $z_{1-\alpha/2}$ je príslušný kvantil normálneho rozdelenia a $s_{\hat{\beta}_j}$ sú štandardné chyby odhadov parametrov.

Na základe Waldových intervalov spoľahlivosti posudzujeme aj významnosť jednotlivých parametrov modelu. Ak príslušný interval obsahuje číslo 0, tak parameter nie je významný na zvolenej hladine významnosti. Z hľadiska interpretácie logistického modelu nie sú tak potrebné parametre a ich intervaly spoľahlivosti, ako sú potrebné pomery šancí (odds ratio) a ich intervaly spoľahlivosti.

SAS EG v procedúre LOGISTIC poskytuje výpočet obidvoch druhov intervalov spoľahlivosti pre parametre logistického modelu: intervaly spoľahlivosti počítané na základe vierohodnostnej funkcie (Profile likelihood)⁵ a aj Waldove intervaly spoľahlivosti.

6. Odhad pomerov šancí a ich intervalov spoľahlivosti

Za predpokladu, že modelovaná premenná je binárna, vysvetlíme si výpočet pomerov šancí (odds ratio) a ich interpretáciu. Označme ako Y binárnu závislú premennú, ktorá nadobúda dve možné hodnoty 0 (neúspech, resp. želaná udalosť nenastala) a 1 (úspech, resp. želaná udalosť nastala). Uvažujme len jednu vysvetľujúcu premennú X_1 . Označme symbolom p podmienenú pravdepodobnosť výskytu želanej udalosti $P(Y=1/X_1)$, teda hodnotu, ktorú modelujeme. Potom $\text{logit}(p)$ zapíšeme v tvare:

⁵ Vzorec pre intervaly spoľahlivosti počítané na základe vierohodnostnej funkcie (*Profile likelihood*) je komplikovaný a tak ho v príspevku neuvádzame.

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) = \log\left(\frac{P(Y=1/X_1)}{P(Y=0/X_1)}\right) = \beta_0 + \beta_1 X_1$$

Šanca je definovaná ako $p/(1-p)$ (angl. výraz je odds). Pomer šancí (OR=odds ratio) je definovaný ako pomer dvoch rôznych šancí. Najčastejšie je to pomer dvoch šancí pri jednotkovej zmene nezávislej premennej X_1 . Logaritmus pomeru šancí pre uvažovaný príklad teda zapíšeme nasledovne:

$$\begin{aligned}\log(OR) &= \log\left(\frac{\text{Šanca}_1}{\text{Šanca}_0}\right) = \log\left(\frac{P(Y=1/X_1=1)/P(Y=0/X_1=1)}{P(Y=1/X_1=0)/P(Y=0/X_1=0)}\right) = \\ &= \log\left(\frac{P(Y=1/X_1=1)}{P(Y=0/X_1=1)}\right) - \log\left(\frac{P(Y=1/X_1=0)}{P(Y=0/X_1=0)}\right) = \\ &= (\beta_0 + \beta_1) - \beta_0 = \beta_1\end{aligned}$$

Hodnotu pomeru šancí OR získame odlogaritmovaním: $OR = \exp(\beta_1)$. V prípade, že sme používali prirodzený logaritmus, tak $OR = e^{\beta_1}$. Pomer šancí nám hovorí, ako sa zmení šanca na úspech (pre modelovanú hodnotu $Y=1$) pri jednotkovej zmene vysvetľujúcej premennej. Napríklad ak $OR=2$, tak šanca na úspech pre $X_1=1$ je dvakrát vyššia, ako keď $X_1=0$.

Ak zmena vysvetľujúcej premennej nie je jednotková, teda z 0 na 1, ale všeobecne z a na b, teda o $c=b-a$, tak OR sa rovná: $OR = \exp(\beta_1)^c$.

Výstup SAS EG pre procedúru LOGISTIC obsahuje okrem odhadov parametrov modelu aj bodové odhady pomerov šancí a ich intervaly spoľahlivosti. Sú viac vhodné na interpretáciu logistického modelu. Z intervalov je možné posúdiť aj významnosť jednotlivých parametrov modelu. Ak príslušný interval spoľahlivosti pre OR obsahuje číslo 1, tak parameter nie je významný⁶.

Vo výskumných prácach hlavne z oblasti medicíny sa často vyskytuje aj pojem risku, resp. rizika (angl. výraz *risk* = RR), ktorý je pre uvažovaný príklad definovaný nasledovne:

$$RR = \frac{P(X_1=1/Y=1)}{P(X_1=1/Y=0)}.$$

7. Hodnotenie kvality logistického modelu

Na posúdenie kvality logistického modelu sa používajú dva druhy mier. Prvá skupina mier je založená na L štatistike a sú to nasledovné ukazovatele:

- Logaritmická transformácia L štatistiky, čiže $-2\log L$. Čím je jej hodnota vyššia, tým je model kvalitnejší, tzn. lepšie odhaduje danú vzorku dát. Má chí-kvadrát rozdelenie a túto štatistiku sme popísali v časti 3.
- Akaikovo informačné kritérium (AIC = Akaike Information Criterion), ktoré predstavuje určitú modifikáciu L štatistiky a výpočtový vzorec má nasledovný tvar:

⁶ V prípade ak sa pomer šancí rovná číslu 1 ($OR=1$), tak sú šance na úspech rovnaké pri rôznych hodnotách vysvetľujúcej premennej. Znamená to, že jednotková zmena premennej X_i nevyvolá zmenu modelovanej podmienenej pravdepodobnosti a tak je táto premenná v modeli nevýznamná.

$$AIC = -2\log L + 2(k + s) \quad (7)$$

- kde s je počet hodnôt (úrovní) modelovanej premennej znížený o číslo 1 (napr. pre binárnu premennú je $s=2-1=1$) a k predstavuje počet vysvetľujúcich premenných v modeli.
- Schwarzovo kritérium, resp. Schwarz-Bayesovo kritérium, (SC = Schwarz Criterion) sa vypočíta podľa vzorca:

$$SC = -2\log L + (k + s)\log(n) \quad (8)$$

kde s je počet hodnôt (úrovní) modelovanej premennej znížený o číslo 1, k predstavuje počet vysvetľujúcich premenných v modeli a n je rozsah vzorky.

Obidve kritéria (AIC aj SC) patria medzi penalizačné miery kvality odhadu regresného modelu. Používajú sa pri porovnaní konkurenčných modelov odhadnutých pre danú vzorku. Nižšie hodnoty týchto mier indikujú vhodnejší model pre dané dáta.

Druhá skupina mier na posúdenie kvality logistického modelu sú miery asociácie, konkrétne Somerovo D (Somer's D), Goodman-Kruskalova gamma (Gamma), Kendallove tau (Tau-a) a štatistika c. Merajú asociáciu medzi predikovanými pravdepodobnosťami pre modelovanú obmenu závislej premennej a skutočnými hodnotami závislej premennej. Čím vyššia je asociácia medzi týmito hodnotami, tým je model kvalitnejší pri predikcii závislej premennej. Miery vychádzajú z párovania pozorovaní na základe hodnôt modelovanej premennej. Zo skutočných (pozorovaných) hodnôt modelovanej premennej a hodnôt predikovaných logistickým modelom môžeme zostaviť asociačnú tabuľku a z nej zistiť nasledovné počty párov pozorovaní:

- n_c - počet zhodných (concordant) párov,
- n_d - počet opačných (discordant) párov,
- n_t - počet ostatných (tied) párov, pre ktorý platí $n_t = t - n_c - n_d$,
- t - celkový počet párov, pre ktorý platí $t = n_c + n_d + n_t$,
- n - počet pozorovaní vo vzorke.

Vzorce pre výpočet jednotlivých mier poradovej asociácie sú:

$$c = (n_c + 0,5(t - n_c - n_d))/t = (n_c + 0,5n_t)/t$$

$$\text{Somers' D} = (n_c - n_d)/t$$

$$\text{Goodman - Kruskal Gamma} = (n_c - n_d)/(n_c + n_d)$$

$$\text{Kendall's Tau - a} = (n_c - n_d)/0,5n(n-1)$$

V praxi sa hlavne používa štatistika c. Model s vyššou hodnotou c štatistiky je kvalitnejší. Je zrejmé, že základom kvality modelu je vysoký počet zhodných párov n_c . V prípade binárnej modelovanej premennej hodnota c štatistiky predstavuje plochu pod ROC krivkou (ROC = Receiver Operating Characteristic), ktorú je možné zobrazit' aj v interaktívnej aplikácii SAS Enterprise Guide v úlohe LOGISTIC.

8. Logistická regresia v systéme SAS

V štatistickom module systému SAS/STAT® je možné použiť viacero procedúr na vykonanie logistickej regresnej analýzy. Sú to procedúry: CATMOD, GENMOD, LOGISTIC

a PROBIT. Každá z týchto procedúr má špeciálne vlastnosti, ktoré sú využiteľné pre konkrétnu aplikáciu.

V aplikácii SAS EG umožňuje v ponukovom režime odhadnúť parametre logistického modelu pre diskretnú závislú premennú metódou maximálnej vierohodnosti len procedúra LOGISTIC. Táto procedúra umožňuje rovnako dobre analyzovať binárnu nominálnu, ale aj poradovú (ordinálnu) závislú premennú. Pre multinomickú ordinálnu závislú premennú procedúra odhaduje logistický model proporcionálnych šancí. Metóda maximálnej vierohodnosti sa vykoná buď Fisherovým alebo Newton-Raphsonovým algoritmom. Linková funkcia môže byť logit, probit alebo aj komplementárna log-log funkcia.

9. Príklad

Na ilustráciu logistickej regresie použijeme jednoduchý príklad o úspešnosti 24 finančných inštitúcií (FI). Premennú, ktorú budeme modelovať (závislá premenná), je úspešnosť inštitúcie (SUCCESS). Ide o binárnu premennú, ktorá nadobúda hodnoty: 1 = veľmi úspešná FI a 2 = málo úspešná FI. Ako nezávislé premenné použijeme (Tabuľka 1):

- veľkosť inštitúcie (SIZE), ktorá nadobúda hodnoty 1 = veľká inštitúcia, 0 = malá inštitúcia a je to binárna premenná;
- určitý finančný ukazovateľ (FP), ktorý je spojitou veličinou.

Tabuľka 1: *Dáta o úspešnosti finančných inštitúcií*
(zdroj dát: [Sharma, 1996])

SUCCESS	SIZE	FP	SUCCESS	SIZE	FP
1	1	0.58	2	1	2.28
1	1	2.80	2	0	1.06
1	1	2.77	2	0	1.08
1	1	3.50	2	0	0.07
1	1	2.67	2	0	0.16
1	1	2.97	2	0	0.70
1	1	2.18	2	0	0.75
1	1	3.24	2	0	1.61
1	1	1.49	2	0	0.34
1	1	2.19	2	0	1.15
1	0	2.70	2	0	0.44
1	0	2.57	2	0	0.86

Úlohy:

- 1) Vytvorte jednoduchý regresný model závislosti úspešnosti finančnej inštitúcie od hodnoty finančného ukazovateľa FP. Výsledky vhodne interpretujte a zobrazte S-krivku pre modelovanú pravdepodobnosť.
- 2) Vytvorte jednoduchý regresný model závislosti úspešnosti finančnej inštitúcie od jej veľkosti SIZE. Výsledky vhodne interpretujte.
- 3) Vytvorte viacnásobný regresný model závislosti úspešnosti finančnej inštitúcie od oboch vysvetľujúcich premenných (FP aj SIZE). Výsledky vhodne interpretujte a porovnajte s ostatnými modelmi.

Riešenie úlohy č. 1:

Úlohou je vytvoriť regresný model závislosti. Pretože závislá premenná SUCCESS je binárna, použijeme model logistickej, a nie lineárnej regresie. Vytvoríme jednoduchý logistický regresný model závislosti úspešnosti finančnej inštitúcie od hodnoty finančného

ukazovateľa FP. V systéme SAS použijeme procedúru LOGISTIC. Ak sú vstupné údaje uložené v SAS dátovom súbore s názvom *FI.sas7bdat* a sú v SAS knižnici WORK, tak kód pre riešenie nášho príkladu je nasledovný:

```
PROC LOGISTIC DATA=FI;
  MODEL SUCCESS = FP / CTABLE PPROB=(0.5);
  OUTPUT OUT=PRED1 P=PHAT;
RUN;
PROC PRINT;
  VAR SUCCESS SIZE PHAT;
RUN;
```

Výsledné odhadnuté podmienené pravdepodobnosti budú uložené do SAS knižnice WORK do dátového súboru s názvom *PRED1.sas7bdat* a to do premennej PHAT a budú aj vytlačené pomocou procedúry PRINT.

Na základe výsledkov môžeme zapísať rovnicu pre bodový odhad logistického modelu č. 1 nasledovne: $\text{logit}(\hat{p}) = \ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = -3,9195 + 2,3782FP$.

Obidva parametre modelu na hladine významnosti $\alpha=0,05$ sú štatisticky významné, pretože p-hodnoty pre Waldove chí-kvadrát testy nadobudli nižšie hodnoty ako číslo $\alpha=0,05$ (pre konštantu (Intercept) je p-hodnota = 0,0087 a pre parameter pri premennej FP je p-hodnota = 0,0045).

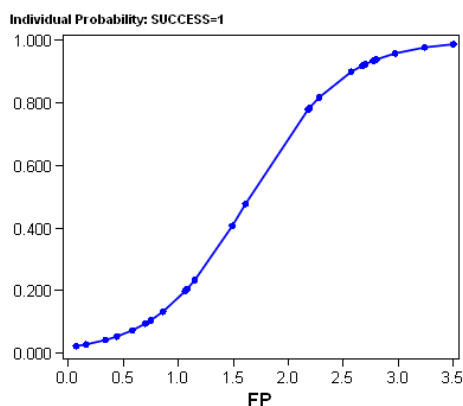
Bodový odhad pomeru šancí pre premennú FP vypočítame takto: $e^{2,3782} = 10,785$ (Tabuľka 3). Výsledný logistický model 1 budeme interpretovať na základe tohto pomeru šancí. Môžeme konštatovať, že šance finančnej inštitúcie stať sa veľmi úspešnou (SUCCESS=1) v porovnaní s možnosťou stať sa málo úspešnou (SUCCESS=2) sa s rastom hodnoty finančného ukazovateľa FP o 1 mernú jednotku v priemere zvyšujú 10,785 krát, resp. sú približne 11 ku 1. Waldov interval spoľahlivosti pre tento pomer šancí je pomerne široký (2,088; 55,718), čo je spôsobené hlavne malým rozsahom vzorky.

Tabuľka 2: Odhad parametrov pre model 1

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.9195	1.495	6.8733	0.0087
FP	1	2.3782	0.8378	8.0573	0.0045

Tabuľka 3: Odhad pomeru šancí pre model 1

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
FP	10.785	2.088	55.718



Obrázok 3: S-krivka vyjadrujúca závislosť podmienenej pravdepodobnosti PHAT od FP

S-krivku pre odhadnutú podmienenú pravdepodobnosť PHAT v závislosti od ukazovateľa FP získame, keď vytvoríme graf závislosti, tzv. Scatter plot (procedúra GPLOT). Výsledná S-krivka (*Obrázok 3*) dokazuje, že modelovaný vzťah má nelineárny charakter. Vidíme, že s rastom finančného ukazovateľa FP pravdepodobnosť inštitúcie, že bude veľmi úspešná, nelineárne rastie.

Riešenie úlohy č. 2:

Opäť vytvoríme jednoduchý logistický model závislosti úspešnosti finančnej inštitúcie od jej veľkosti. Vysvetľujúca premenná SIZE je binárna. SAS kód je analogický kódu z predchádzajúcej úlohy.

Na základe výsledkov získavame bodový odhad logistického modelu č. 2 v tvare:

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = -1,7047 + 4,0073\text{SIZE}$$

Obidva parametre sú na hladine významnosti $\alpha = 0,05$ štatisticky významné, lebo ich p-hodnoty sú nízke (*Tabuľka 4*). Logistický model interpretujeme pomocou šancí. V našom prípade sú šance finančnej inštitúcie stať sa veľmi úspešnou (SUCCESS=1) pre veľké inštitúcie (SIZE=1) 55 krát väčšie ($e^{4,0073} = 54,999$) ako pre malé finančné inštitúcie (SIZE=0), resp. sú 55 ku 1 (*Tabuľka 5*).

Tabuľka 4: Odhad parametrov pre model 2

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.7047	0.7687	4.9181	0.0266
SIZE	1	4.0073	1.3003	9.4971	0.0021

Tabuľka 5: Odhad pomeru šancí pre model 2

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
SIZE	54.999	4.3	703.406

Bodové odhady podmienených pravdepodobností pre jednotlivé štatistické jednotky (FI) pri rôznych úrovniach nezávislej premennej SIZE vypočítame podľa vzorca:

$$\hat{p} = \frac{1}{1 + e^{-(-1,7047 + 4,0073\text{SIZE})}} \quad .$$

Pretože premenná SIZE je binárna, výsledkom sú len dve hodnoty podmienenej pravdepodobnosti. Ak SIZE = 0, tak je táto pravdepodobnosť 0,154. Znamená to, že pravdepodobnosť malej organizácie stať sa veľmi úspešnou je odhadovaná modelom len na 15,4%. Pre SIZE = 1 je táto pravdepodobnosť 0,909. Pravdepodobnosť veľkej organizácie stať sa veľmi úspešnou je odhadovaná modelom až na 90,9%.

Riešenie úlohy č. 3:

Úlohou je vytvoriť viacnásobný logistický model závislosti s dvomi vysvetľujúcimi premennými FP a SIZE. SAS kód pre získanie výsledkov a ich uloženie do SAS súboru do knižnice WORK s názvom PRED2.sas7bdat je:

```
PROC LOGISTIC DATA=FI;
  MODEL SUCCESS = SIZE FP / CTABLE PPROB=(0.5);
  OUTPUT OUT=PRED2 P=PHAT;
RUN;
```

Výsledný bodový odhad logistického modelu číslo 3 je (*Tabuľka 6*):

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = -4,4449 + 1,9244FP + 3,0552SIZE$$

Modelovaný logaritmus šancí (logit) inštitúcie stať sa veľmi úspešnou sa bude pri vzraste finančného ukazovateľa FP o 1 mernú jednotku zvyšovať v priemere o 1,9244 (za predpokladu, že sa nebude meniť jej veľkosť SIZE). Druhý parameter udáva, že ak sa zmení veľkosť inštitúcie z úrovne „malá“ FI, t.j. SIZE=0, na úroveň „veľká“ FI, t.j. SIZE=1, tak vzrastie v priemere logit o hodnotu 3,0552 (za predpokladu, že sa nebude meniť finančný ukazovateľ FP).

Model 3 je ako celok štatisticky významný na hladine významnosti $\alpha=0,05$. Na tejto hladine nie sú však významné všetky jeho parametre, konkrétne vysokú p-hodnotu pre Waldov chí-kvadrát test má parameter pri premennej SIZE (p-hodnota=0,0559). Táto skutočnosť je zrejmá aj z Waldovho 95%-ného intervalu spoľahlivosti pre pomer šancí (Tabuľka 8). Tento interval obsahuje číslo 1 a je veľmi široký. Podľa výsledkov 95%-ných intervalov spoľahlivosti pre pomery šancí vypočítaných na základe vierohodnostnej (Likelihood) štatistiky (Tabuľka 7), by sme parameter modelu 3 pri SIZE nezamietli (interval obsahuje len hodnoty väčšie ako 1), ale horná hranica tohto intervalu je v nekonečne. Rôzne výsledky a veľká šírka intervalov spoľahlivosti pre parametre sú spôsobené malým rozsahom vzorky. Ako sme spomínali už v teoretickej časti, pre malé vzorky sú vhodnejšie výpočty na základe vierohodnostného pomeru a nie Waldovej štatistiky. Waldove odhady sú pre malé vzorky často vychýlené.

Tabuľka 6: Parametre a ich testy významnosti pre model 3

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-4.4449	1.8432	5.8156	0.0159
FP		1	1.9244	0.9116	4.4568	0.0348
SIZE	0	1	3.0552	1.5981	3.6548	0.0559

Tabuľka 7: Pomery šancí pre model 3 (Likelihood)

Profile Likelihood Confidence Interval for Adjusted Odds Ratios				
Effect	Unit	Estimate	95% Confidence Limits	
FP	1.0000	6.851	1.448	66.259
SIZE	1.0000	21.225	1.191	>999.999

Tabuľka 8: Pomery šancí pre model 3 (Wald)

Odds Ratio Estimates		
Effect	Point Estimate	95% Wald Confidence Limits
FP	6.851	1.148 40.896
SIZE	21.225	0.926 486.565

Na záver porovnáme miery kvality pre všetky 3 modely (Tabuľka 9). Z tabuľky vidíme, že model 3 má najnižšie hodnoty kritérií AIC a SC a najvyššie hodnoty štatistiky c a percenta zhodných párov, preto by bol najkvalitnejší. Ďalej však konštatujeme, že samostatne sú obe premenné FP a SIZE významné pri modelovaní podmienenej pravdepodobnosti, ale pri zaradení oboch premenných naraz do logistického modelu, premenná SIZE je na hladine významnosti $\alpha=0,05$ nevýznamná. Tieto výsledky sú ovplyvnené malým rozsahom vzorky dát ($n = 24$). Uvedený príklad bol použitý z metodických dôvodov. V praxi je potrebné používať väčšie reprezentatívne výberové súbory.

Tabuľka 9: Porovnanie mier kvality a asociácie pre odhadnuté logistické modely

Criterion	Model 0	Model 1 (len FP)	Model 2 (len SIZE)	Model 3 (FP a SIZE)
	Intercept Only	Intercept and Covariates	Intercept and Covariates	Intercept and Covariates
AIC	35.271	20.115	21.864	17.789
SC	36.449	22.471	24.221	21.323
-2 Log L	33.271	16.115	17.864	11.789
c		0.917	0.875	0.958
Percent Concordant		91.7	76.4	95.8

10. Záver

V procese modelovania dát môžeme použiť množstvo rôznych metód, algoritmov a techník. Výber druhu modelu závisí od cieľa, pre ktorý je model určený a tiež od typu závislej premennej. Proces tvorby modelu pozostáva z rady krokov, ktoré treba pre správne fungovanie modelu v praxi dodržať.

Logistická regresia je pri správnom použití veľmi účinná, priamočiara a nehrozí pri nej prílišné prispôsobenie sa modelu dátam (overfitting). Je vynikajúcou metódou pre nájdenie nelineárnych závislostí v dátach pri minimálnej chybe. Využíva kategoriálne a spojité premenné na predpoveď výsledku, ktorý má podobu kategórií. Skúsenosti z praxe hovoria, že keď si dáme prácu s dôslednou prípravou a transformáciou premenných, tak výsledný model spracovaný pomocou logistickej regresie prekonáva všetky iné metódy a je výkonný.

11. Literatúra

HEBÁK, P. A KOLEKTÍV: Vícerozmerné statistické metody (3). Informatorium. Praha 2005. ISBN 80-7333-039-3

KLEINBAUM, DAVID G.: Logistic regression. A Self-Learning Text. Springer. 1994

Luha J.: Analýza nominálnych a ordinálnych znakov. *EKOMSTAT 2000*, Trenčianske Teplice, SŠDS. ISBN 80-88946-04-2

Luha J.: Viacrozmerné štatistické metódy analýzy kvalitatívnych znakov. *EKOMSTAT 2005*, Štatistické metódy v praxi. SŠDS Trenčianske Teplice 22. – 27. 5. 2005. ISBN 80-88946-43-3

ŘEZÁNKOVÁ, H.: Analýza kategoriálních dat. Nakladatelství Oeconomica. VŠE Praha. 2005. ISBN 80-245-0926-01

STANKOVIČOVÁ, IVETA: Logická regresia a hĺbková analýza dát (Data Mining). In: Štatistické metódy v praxi. Bratislava: SŠDS 2002. ISBN 80-88946-19-0

STANKOVIČOVÁ, IVETA: Grafy interakcií. Zobrazovanie výsledkov logistických modelov. In: Výpočtová štatistika : SŠDS, 2003. Str. 105-109. ISBN 80-88946-29-8

SHARMA, SUBHASH: Applied Multivariate Techniques. New York, John Wiley & Sons, Inc., 1996.

Adresa autora:

Ing. Iveta Stankovičová, PhD.
Katedra informačných systémov
Fakulta managementu UK v Bratislave
Odbojárov 10, P. O. Box 95
820 05 Bratislava 25
iveta.stankovicova@fm.uniba.sk

Príspevok bol spracovaný v rámci riešenia grantovej úlohy VEGA 1/2631/05 „Analýza možností aplikácie viacrozmerných štatistických metód na skúmanie ekonomických výsledkov na príklade priemyslu SR prípadne iných oblastí ekonomiky“.

Dvojkroková zhluková analýza – automatické zistenie počtu zhlukov

RNDr. Samuel Koróny*

Abstract: The paper provides introduction and some application results of two step cluster analysis. It is standard part of SPSS 11.5 and later versions. It seems that it is one of the best available cluster algorithms.

Keywords: Cluster analysis, Statistical analysis.

1. Úvod

Od začiatku 90.-tych rokov 20. storočia sa intenzívne rozvíjali metódy na zhukovanie veľkých dátových súborov. Bolo to spôsobené aj boomom v oblasti data miningu. Dvojkroková zhluková analýza ako jeden z výsledkov výskumov sa zaradila medzi relatívne osvedčené zhlukovacie postupy.

2. Požiadavky

Ideálna zhluková metóda by mala spĺňať nasledovné kritériá (Hebák 2005): minimálna náročnosť (strojový čas, pamäť...), nezávislosť na poradí vstupu objektu do analýzy, možnosť zhodnotenia a interpretácie nájdených zhlukov, dimenzionalita (efektívnosť aj v prípade veľkého počtu premenných), vyriešenie odľahlých objektov, schopnosť pracovať s rôznymi typmi premenných (spojité aj kategoriálne spolu).

3. Popis dvojkrokovej zhlukovej analýzy

Dvojkroková zhluková analýza patrí medzi metódy používajúce postupy data miningu a data squashingu (drvenia dát). Metódy drvenia dát vypočítajú zo skupiny objektov určité charakteristiky, ktoré sú použité pre ďalšie zhukovanie namiesto pôvodných údajov. Príkladom je tzv. CF (cluster feature) štatistika, ktorá je základom algoritmu BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies, (vid' Zhang 1996)). Objekty sa zaradia do podzhlukov, charakterizovaných CF a potom sa ďalej zhukujú pomocou tradičnej hierarchickej zhlukovej metódy. CF štatistika obsahuje tri údaje: počet objektov,

* RNDr. Samuel Koróny, Ústav vedy a výskumu UMB, Banská Bystrica

súčet hodnôt príslušnej premennej a súčet hodnôt druhých mocnín premennej. Metóda BIRCH je použitá aj v prípade dvojkrokovej zhlukovej analýzy, ktorá vie pracovať so spojitými aj kategorickými premennými. Vyžaduje len jeden výpočtový prechod cez dáta. Má dva stupne: najprv je predbežné zhlukovanie objektov (riadkov) do mnohých malých predbežných zhlukov, potom nasleduje spojenie predbežných zhlukov do požadovaného počtu zhlukov. Dokáže aj nájsť optimálny počet zhlukov. Je súčasťou systému SPSS od verzie 11.5.

3.1 Prvý krok

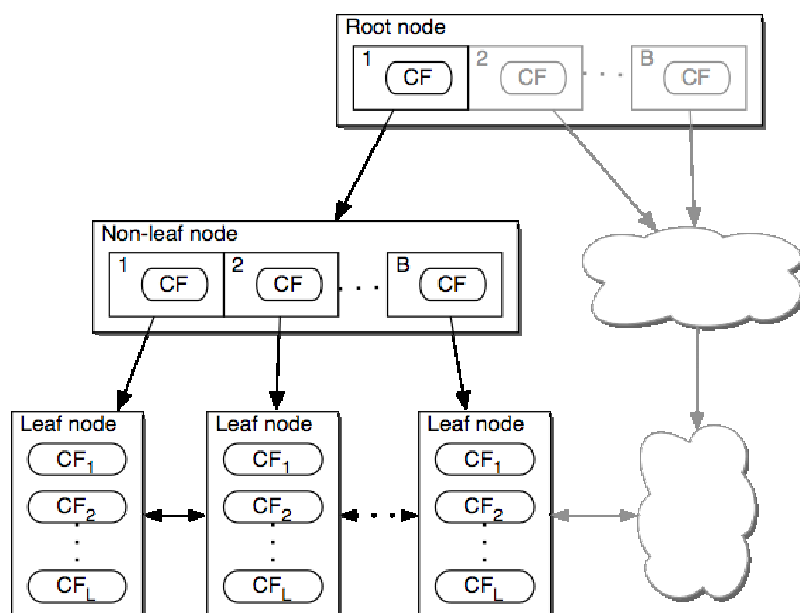
V prvom kroku sa používa sekvenčný zhlukový postup. Prechádzajú sa postupne dáta a rozhoduje sa, či daný objekt má byť pripojený k už vytvorenému podzhluku alebo má byť zdrojom pre nový podzhluk na základe kritéria vzdialenosti. Počet podzhlukov je podstatne menší ako počet objektov. Je vhodné, aby boli objekty náhodne usporiadané. Algoritmus postupne vytvára CF strom, ktorý obsahuje listy a nelistové uzly v niekoľkých úrovniach. Každý uzol má niekoľko vstupov. Samotný CF strom je určitá forma viacúrovňovej kompresie dát (viď obr.1). Listy stromu sú konečné podzhluky. Nelistový uzol má maximálne B potomkov a list má najviac L vstupov. Každý vstup obsahuje tri spomenuté CF charakteristiky a početnosti kategórií kategorických premenných. Pri vkladaní objektu sa postupuje nasledovne: ak je objekt v rámci prahovej vzdialenosti od najbližšieho podzhluku, potom vstúpi do neho. Pritom sa aktualizuje príslušná CF charakteristika. Ak je objekt ďalej, tak sám vytvára nový podzhluk. Ak už nie je možné vytvoriť ďalší podzhluk, tak sa podzhluk rozdelí na dva na základe najvzdialenejších vstupov. Ostatné vstupy sa prerozdedia podľa kritéria vzdialenosti. Výsledkom prvého kroku je nová dátová matica, kde v riadkoch sú charakteristiky jednotlivých podzhlukov.

Štandardne nastavené parametre v systéme SPSS sú:

Maximálny počet úrovní stromu	3
Maximálny počet vstupov do uzla	8
Maximálny počet podzhlukov	512
Prahová vzdialenosť	0

Prahová vzdialenosť sa nedá meniť užívateľom, začína pri hodnote 0 a zmení sa vždy na najmenšiu vzdialenosť medzi CF objektami pri potrebe zmeniť CF strom.

Algoritmus dvojkrokovej zhlukovej analýzy bol testovaný na mnohých simulovaných aj reálnych dátach a bolo zistené, že maximálny počet podzhlukov 512 by mal stačiť na presné a relatívne rýchlo dosiahnuteľné výsledky. Meniť nastavené parametre sa odporúča len v prípade znalostí o lepších výsledkoch.



Obr. 1 Schematický náčrt CF stromu

3.2 Druhý krok

V druhom kroku sú podzhluky z prvého kroku spájané do požadovaného počtu zhlukov tradičnou hierarchickou metódou.

V obidvoch krokoch sa používajú miery vzdialenosti, pričom v SPSS sú implementované dve: euklidovská a vierohodnostná. Euklidovská sa môže použiť len v prípade spojitých premenných.

Označenie vo vzorcoch pre dvojkrovovú zhlukovú analýzu :

K^A je celkový počet použitých spojitých premenných,

K^B je celkový počet použitých kategorických premenných,

L_k je počet kategórií k -tej kategorickej premennej,

R_k je rozsah k -tej spojitej premennej,

N je celkový počet údajov,

N_k je celkový počet údajov v k -tom zhluku,

$\hat{\sigma}_k^2$ odhad rozptylu k -tej spojitej premennej,

$\hat{\sigma}_{jk}^2$ odhad rozptylu k -tej spojitej premennej v j -tom zhluku

N_{jkl} je počet hodnôt l -tej kategórie k -tej kategorickej premennej v j -om zhluku

$d(j, s)$ je vzdialenosť medzi j -tým a s -tým zhlukom

$\langle j, s \rangle$ index pre zhluk vzniknutý spojením j -teho a s -tého zhluku

3.3 Miera vzdialenosti

Vzdialenosť log-likelihood vie pracovať so spojitými aj kategorickými premennými. Je založená na pravdepodobnosti. Vzdialenosť medzi dvomi zhlukmi závisí od poklesu log-likelihoodu pri spojení do jedného zhuku. Pri výpočte log-likelihoodu sa predpokladá normálne rozdelenie pre spojité premenné a multinomické rozdelenie pre kategorické premenné. Tiež sa predpokladá vzájomná nezávislosť premenných a objektov. Vzdialenosť medzi zhlukmi j a s je

$$d(j, s) = \xi_j + \xi_s - \xi_{\langle j, s \rangle}, \quad (1)$$

kde

$$\xi_v = -N_v \left(\sum_{k=1}^{K^A} \frac{1}{2} \log(\hat{\sigma}_k^2 + \hat{\sigma}_{vk}^2) \right) + \sum_{k=1}^{K^B} \hat{E}_{vk}, \quad (2)$$

a

$$\hat{E}_{vk} = - \sum_{l=1}^{L_k} \frac{N_{vkl}}{N_v} \ln \frac{N_{vkl}}{N_v}. \quad (3)$$

3.4 Počet zhlukov

Pre určenie počtu zhlukov sa používajú informačné kritériá AIC a BIC. V prvom kroku sa vypočíta kritérium pre každý počet zhlukov a urobí sa prvý odhad počtu zhlukov. V druhom kroku sa počiatočný odhad spresní nájdením najväčšieho nárastu vo vzdialenosti medzi dvomi najbližšími zhlukmi v každom hierarchickom zhlučovacom štádiu. Vzťah pre bayesovské informačné kritérium BIC pre J zhlukov je:

$$BIC(J) = -2 \sum_{j=1}^J \xi_j + m_J \ln(N),$$

Podobný vzťah platí pre Akaikeho informačné kritérium:

$$AIC(J) = -2 \sum_{j=1}^J \xi_j + 2m_J,$$

kde

$$m_J = J(2K^A + \sum_{k=1}^{K^B} (L_k - 1)).$$

4. Aplikácie

Ako ukážku tu uvádzame výsledky aplikovania dvokrokovej zhlukovej analýzy na reálne údaje z oblasti predaja 2 440 pozemkov, ktoré sú súčasťou systému SPSS v súbore Home sales. Uvedený súbor obsahuje nasledovné spojité premenné: odhad hodnoty pozemku, odhad zhodnotenia pozemku, celková odhadovaná trhovú hodnotu, skutočná trhovú cenu a poloha pozemku. Okrem toho bola do analýzy zahrnutá kategorická premenná poloha (neighborhood), ktorá má 7 hodnôt podľa oblasti umiestnenia. Pre celkový prehľad sú najprv uvedené základné štatistické charakteristiky súboru. Spojité premenné obsahujú počet pozorovaní, aritmetický priemer, smerodajnú odchýlku, minimum a maximum (tab.1). Kategorická premenná má frekvenčnú tabuľku (tab. 2).

Tab. 1 Opisná štatistika spojitých premenných

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Appraised Land Value	2440	1800	93200	17041.04	11174.716
Appraised Value of Improvements	2440	713	99954	40980.58	23382.739
Total Appraised Value	2440	119	99989	48431.45	23816.676
Sale Price	2440	5500	445000	74738.54	49260.688
Valid N (listwise)	2440				

Tab. 2 Frekvenčná tabuľka kategorickej premennej Neighborhood

Neighborhood

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid A	42	1.7	1.7	1.7
B	319	13.1	13.1	14.8
C	258	10.6	10.6	25.4
D	467	19.1	19.1	44.5
E	500	20.5	20.5	65.0
F	372	15.2	15.2	80.2
G	482	19.8	19.8	100.0
Total	2440	100.0	100.0	

Pre zníženie negatívneho vplyvu korelácií nebola do zhlukovej analýzy zahrnutá hodnota pozemku. Výstup dvojškrokového zhlukového analýzy začína časťou auto-clustering, ktorá obsahuje hodnoty zvoleného kritéria na automatické nájdenie počtu zhlukov (tab.3).

V systéme SPSS sa dá aj fixne zvoliť počet zhlukov a vtedy sa rozdelí súbor na zhluky objektov na základe zvoleného typu vzdialenosti. V našom prípade zmiešaných premenných to môže byť vzhľadom na prítomnosť kategorickej premennej len log-likelihood vzdialenosť. Pri určení počtu zhlukov sa za optimálny považuje prípad, keď hodnota v pravom stĺpci nadobúda lokálne maximum. Ide o podiel mier vzdialeností po sebe idúcich zhlukových zostáv. Pomer vzdialeností pre 5 zhlukov a 4 zhluky je 1,742 a preto hľadaný optimálny počet zhlukov podľa daného kritéria je 5. Základný prehľad riešenia je v tab. 4.

Tab. 3 Hodnoty kritérií pre rôzne počty zhlukov

Auto-Clustering				
Number of Clusters	Schwarz's Bayesian Criterion (BIC)	BIC Change(a)	Ratio of BIC Changes(b)	Ratio of Distance Measures(c)
1	14056.772			
2	9756.426	-4300.346	1.000	1.919
3	7559.895	-2196.531	.511	1.518
4	6144.573	-1415.322	.329	1.230
5	5011.852	-1132.721	.263	1.742
6	4401.543	-610.309	.142	1.128
7	3870.948	-530.595	.123	1.460
8	3536.932	-334.015	.078	1.597
9	3362.852	-174.080	.040	1.308
10	3251.854	-110.998	.026	1.110
11	3161.098	-90.756	.021	1.277
12	3110.352	-50.747	.012	1.046
13	3065.890	-44.462	.010	1.235
14	3047.704	-18.186	.004	1.021
15	3031.858	-15.846	.004	1.079

a The changes are from the previous number of clusters in the table.

b The ratios of changes are relative to the change for the two cluster solution.

c The ratios of distance measures are based on the current number of clusters against the previous number of clusters.

Tab. 4 Počet objektov v nájdených zhlukoch

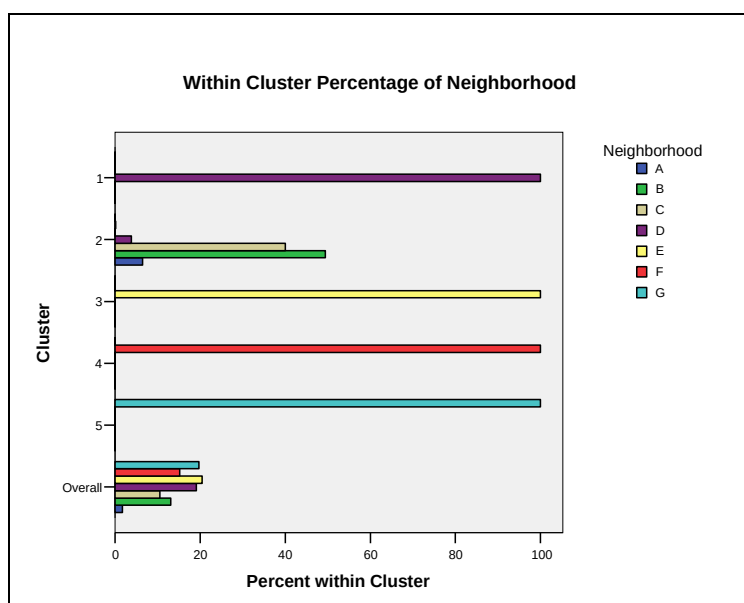
Cluster Distribution				
		N	% of Combined	% of Total
Cluster	1	442	18.1%	18.1%
	2	645	26.4%	26.4%
	3	500	20.5%	20.5%
	4	371	15.2%	15.2%
	5	482	19.8%	19.8%
	Combined	2440	100.0%	100.0%
Total		2440		100.0%

V tab. 5 sú uvedené základné štatistické charakteristiky premenných podľa nájdených zhlukov. Ak je súbor dobre rozdelený, tak daná tabuľka by mala stačiť na interpretáciu zhlukov.

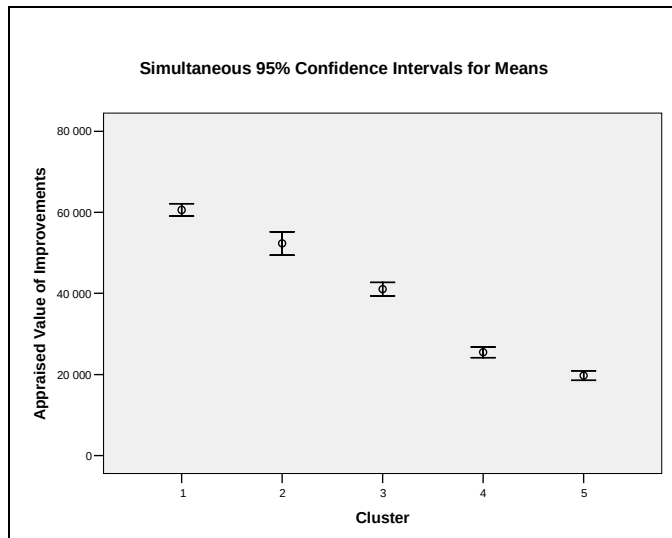
Tab. 5 Základné štatistické charakteristiky spojitých premenných po zhlukoch

		Centroids					
		Appraised Value of Improvements		Total Appraised Value		Sale Price	
		Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
Cluster	1	60575.30	12717.569	72122.48	19077.381	85039.58	22442.996
	2	52301.04	28351.352	48054.14	27933.390	120879.86	64081.540
	3	41047.63	14262.925	53577.05	15936.212	62966.92	25028.825
	4	25471.30	9704.754	39720.49	11244.266	49789.61	16737.824
	5	19731.33	10088.189	28578.56	11841.255	34961.87	17814.035
	Combined	40980.58	23382.739	48431.45	23816.676	74738.54	49260.688

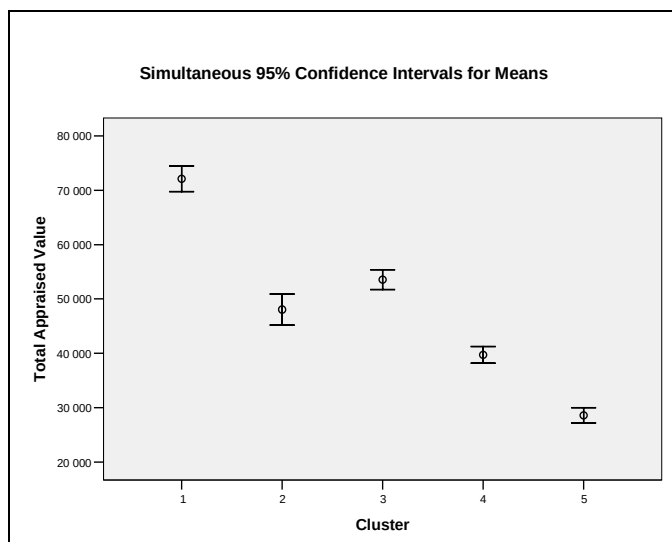
Pre lepšiu orientáciu systém SPSS ponúka výsledky dvoj krokovej zhlukovej analýzy aj prostredníctvom množstva grafov. Prvý z nich je stĺpcový graf podielu jednotlivých kategórií hodnôt po zhlukoch (graf 1). Z grafu je vidieť, že zhluky sú z veľkej miery determinované práve polohou pozemkov. Výnimkou je zhluk 2, ktorý je spoločne obsadený všetkými kategóriami, pričom prevláda kategória B. Grafy 2, 3 a 4 ponúkajú vcelku jednoznačnú interpretáciu zhlukov na základe simultánnych konfidenčných intervalov aritmetických priemerov spojitých premenných: odhad zhodnotenia pozemku, jeho celková odhadovaná trhovú hodnotu a skutočná predajná cena.



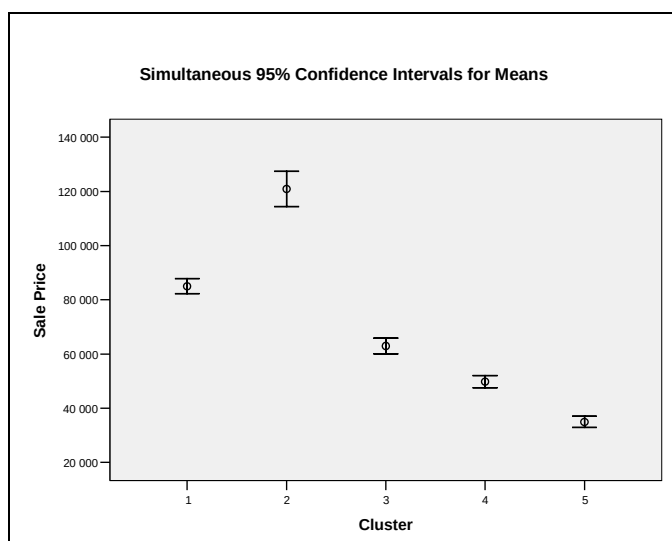
Graf 1 Percentuálny podiel hodnôt premennej neighborhood po zhlukoch



Graf 2 Simultánne 95 % konfidenčné intervaly premennej zlepšenia hodnoty pozemkov



Graf 3 Simultánne 95 % konfidenčné intervaly premennej celkovej hodnoty pozemkov



Graf 4 Simultánne 95 % konfidenčné intervaly premennej predajnej ceny

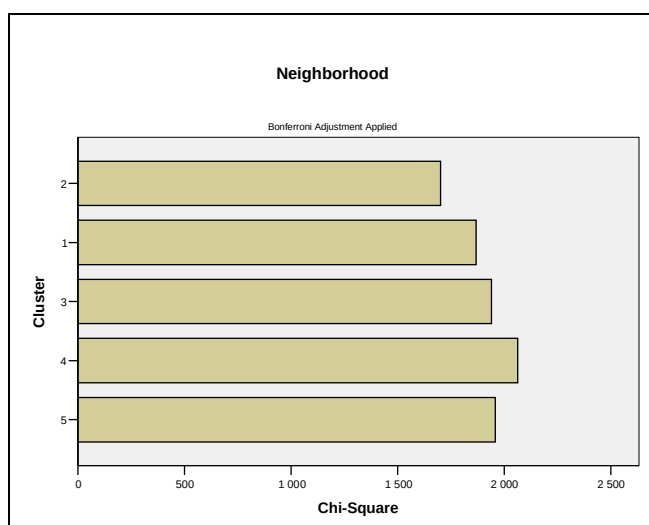
Interpretácia zhlukov

Na základe tabuliek a grafov je možné interpretovať jednotlivé zhluky:

1. zhluk charakterizujú pozemky z oblasti D, ktoré majú najvyššie priemerné zhodnotenie pozemku, tiež najvyššiu odhadovanú hodnotu pozemku a druhú najvyššiu predajnú cenu,
2. zhluk má zastúpenie pozemkov zo všetkých oblastí (prevláda pritom oblasť B), má druhé najvyššie priemerné zhodnotenie pozemku, tretiu najvyššiu odhadovanú hodnotu pozemku a najvyššiu predajnú cenu,
3. zhluk obsahuje len pozemky z lokality E, ktoré majú stredne veľké priemerné hodnoty spojitých premenných,
4. zhluk obsahuje len pozemky z lokality F, ktoré majú druhé najnižšie priemerné hodnoty spojitých premenných,
5. zhluk obsahuje len pozemky z lokality G, ktoré majú celkovo najnižšie priemerné hodnoty spojitých premenných.

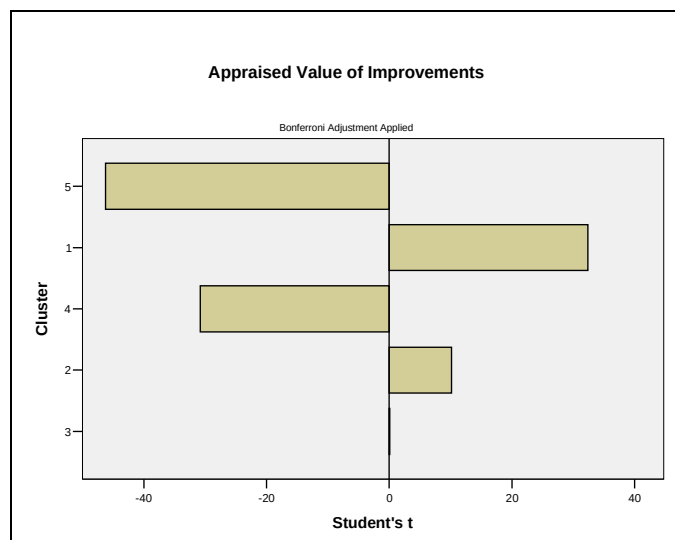
Overenie vplyvu jednotlivých premenných na formovanie zhlukov

Pre kategorické premenné systém SPSS vypočíta hodnotu chí kvadrát, ktorá porovnáva rozdelenie hodnôt premennej v zhluke voči celkovému rozdeleniu hodnôt. Graf 5 zobrazuje chí kvadrát štatistiku pre premennú neighborhood. Kritickú hodnotu nie je vidieť na grafe, pretože splýva s vertikálnou osou. Je to spôsobené vysokou signifikantnosťou rozdelenia hodnôt danej premennej.

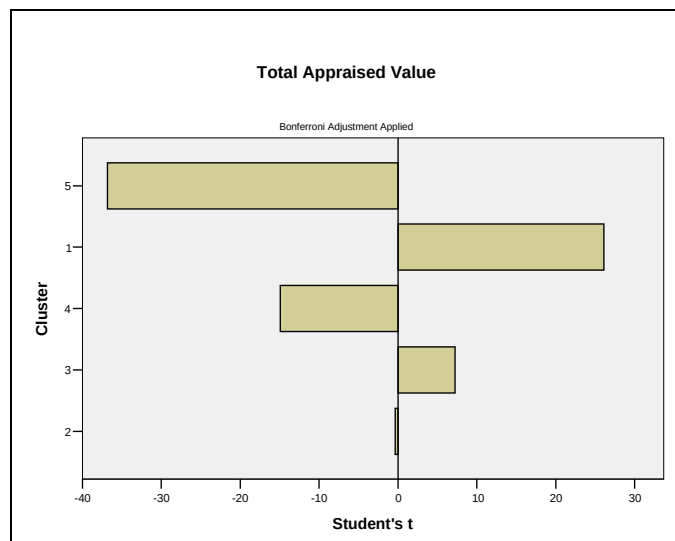


Graf 5 Vplyv premennej neighborhood („poloha“) na formovanie zhlukov

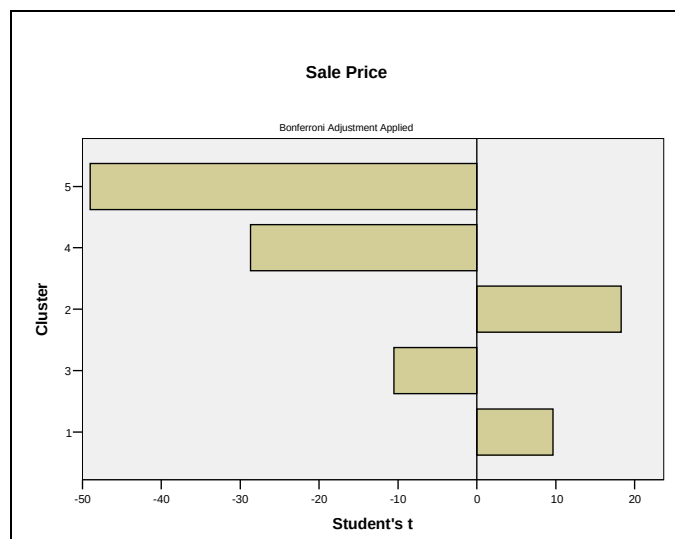
Pre spojité premenné je k dispozícii podobný graf založený na t - štatistike. V ňom sa porovnáva aritmetický priemer premennej v zhluku voči celkovému priemeru. Na grafoch 6, 7 a 8 sú uvedené výsledky pre sledované spojité premenné: odhad zhodnotenia pozemku, jeho celková odhadovaná hodnota, skutočná predajná cena pozemku. Opäť je kritická hodnota nezobrazená, lebo rozdiely sú silne signifikantné.



Graf 6 Vplyv odhadu zhodnotenia pozemku na formovanie zhlukov



Graf 7 Vplyv odhadu celkovej hodnoty pozemku na formovanie zhlukov



Graf 8 Vplyv predajnej ceny pozemku na formovanie zhlukov

Upozornenie

Uvedené „testovanie“ významnosti po zhlukoch v prípade obidvoch typov premenných je len orientačné. Nejde o celkom korektné použitie daných testov. Samozrejme toto platí často v zhlukovej analýze.

5. Záver

Doterajšie skúsenosti s použitím dvoj krokovej zhlukovej analýzy ukazujú, že ide o skutočne veľmi silnú a spoľahlivú zhlukovú metódu, ktorá určite nájde svoje miesto v štandardnej výbave štatistických metód. To samozrejme závisí aj od softwarových firiem.

Použitá literatúra

SPSS 13.0 USER'S GUIDE. Chicago : SPSS Inc., 2004. ISBN 0-13-185723-1

HEBÁK, P. A KOL. 2005. Vícerozměrné statistické metody (3). Praha : Informatorium, 2005. ISBN 80-7333-039-3

KORÓNY, S. 2006. Zhluková analýza produkčných ukazovateľov stavebných podnikov právnej formy s. r. o. In: Forum Statisticum Slovaca. Roč.2, č.1, 2006, s. 56-60. ISSN 1336-7420

CHIU, T. ET AL. 2001. A Robust and Scalable Clustering Algorithm for Mixed Type Attributes in Large Databases Environment. Proceedings of the ACM SIGKDD Conference on knowledge discovery and data mining, p. 263-268, Montreal

ZHANG, T. ET AL. 1996. BIRCH: An Efficient Data Clustering Metod for Very Large Databases. In: Proceedings of the ACM SIGMOD Conference on Management of Data, p. 103-114, Montreal

Kontakt na autora:

RNDr. Samuel Koróny

Ústav vedy a výskumu UMB

Cesta na amfiteáter 1

974 01 Banská Bystrica

Email: samuel.korony@umb.sk

Makroekonomický vývoj SR do apríla 2007

Jozef Chajdiak

Abstract

The article contains graphical presentation of macroeconomical development SR

Celkovo možno náš makroekonomický vývoj charakterizovať ako ešte stále zlý stav s prevažne dobrou tendenciou vývoja.

Vo vývoji HDP s stálych cenách 2000 na obr. 1 môžeme konštatovať dve etapy: do roku 1998 a po roku 1999. Úroveň konca druhej etapy (8 percentný rast v roku 2006) sa približuje začiatku prvej etapy. Na obr. 2 je znázornený analogický rast HDP v bežných cenách. Keďže jeho objem je formovaný okrem zmeny naturálnych objemov produkcie aj zmenou cien, sú hodnoty indexov trochu vyššie a variabilnejšie.

Hospodárenie so štátnym rozpočtom je zobrazené na obr. 3 až obr. 5. V posledných nekoľko mesiacov sa schodok zväčšuje. V poslednom roku sa zlepšila veľkosť schodku štátneho rozpočtu.

Na obr. 6 je vývoj medziročnej inflácie. Ku koncu obdobia možno sledovať zlepšovanie vývoja.

Kurz slovenskej koruny voči euru (obr. 7) a kurz slovenskej koruny voči českej korune (obr. 8) sa v súčasnosti spevňuje.

Mzdový vývoj je na obr. 9 a obr. 10. Vidíme, že oproti roku 1989 (s platmi pod 3000 korún sa v súčasnosti už dostávame na úroveň mierne nad dvadsaťtisíc korún. V reálnej mzde sa už podarilo prekročiť mzdovú úroveň roku 1989.

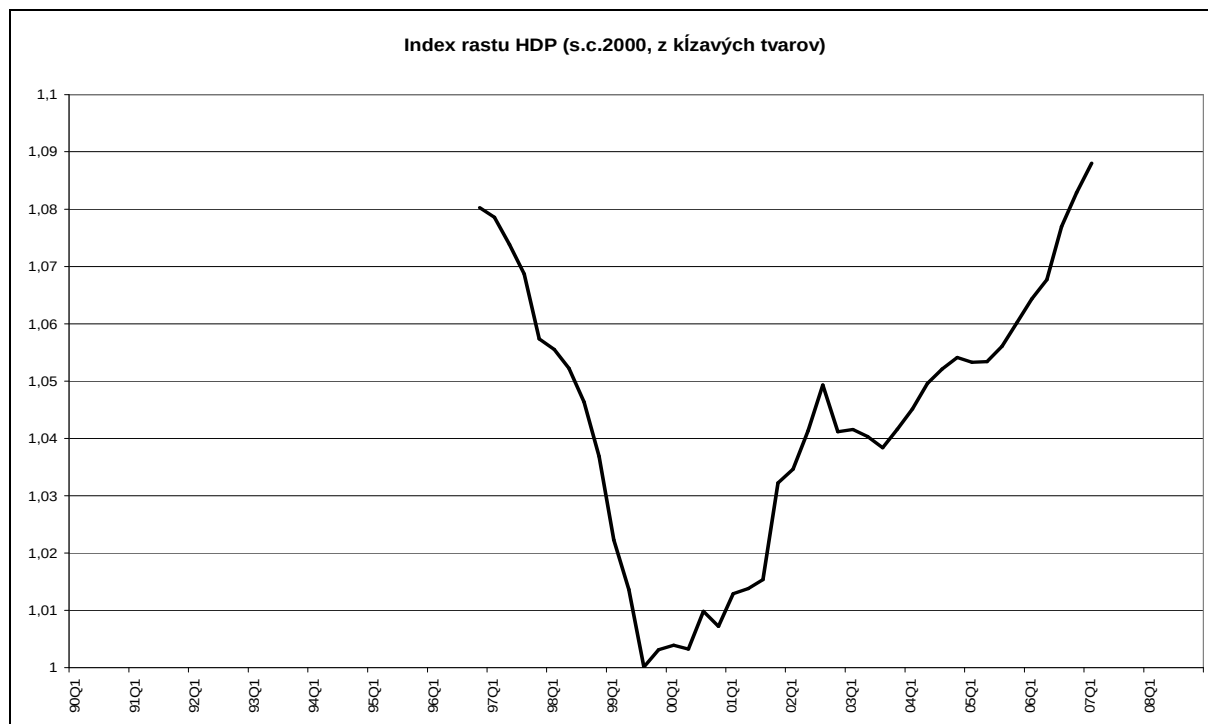
Na obr. 11 je znázornený vývoj celkového počtu nezamestnaných. Oproti vrcholom v rokoch 2000 až 2002 ich počet postupne klesá a dostáva sa pod úroveň 300 tisíc nezamestnaných.

Vývoj zahraničného obchodu je na obr. 12 a obr. 13. V konce obdobia sa zdá, že začala etapa zlepšovania salda zahraničného obchodu.

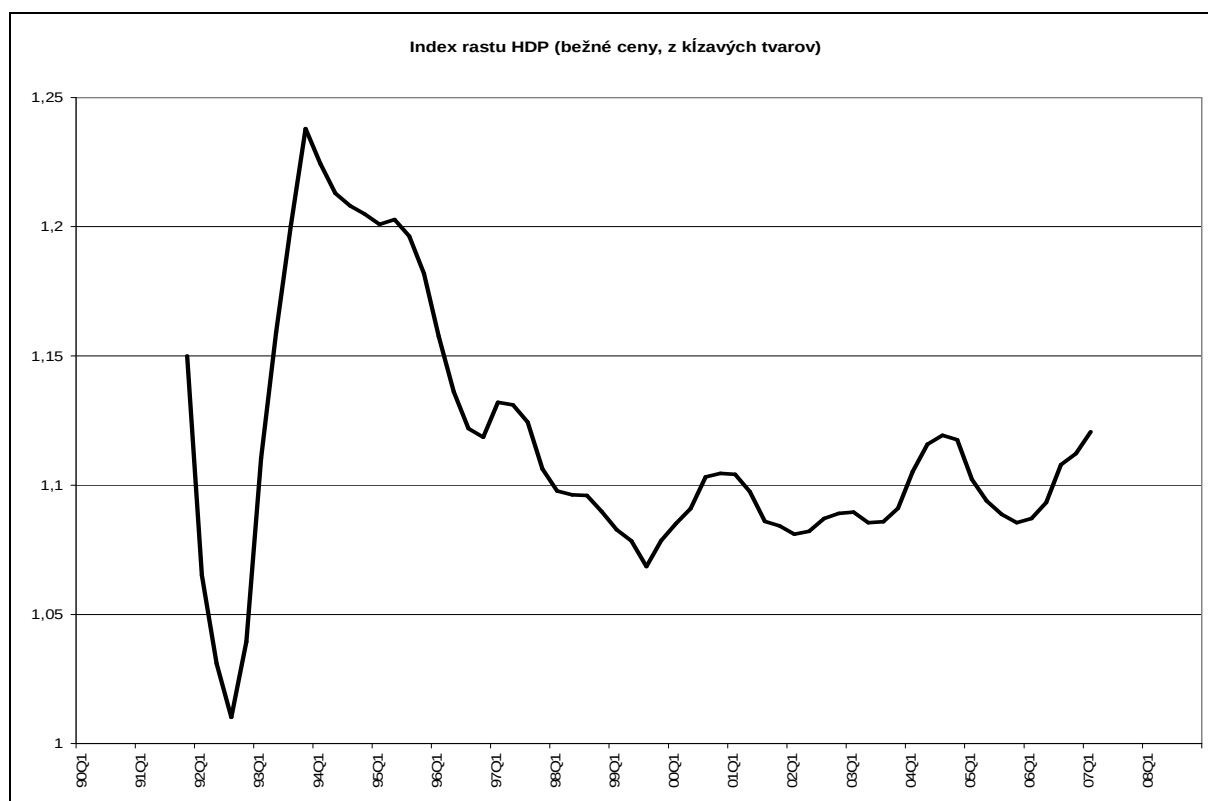
Vývoj kľúčových ročných objemov príjmov štátneho rozpočtu z dane z pridanej hodnoty v prepočte na objem HDP je na (obr. 14). V posledných obdobiach tento podiel klesá, čo znamená, že objem HDP rastie a daň z pridanej hodnoty stagnuje.

Na obr. 15 je vývoj podielu dlhu (maastrichtský) ústrednej vlády na HDP v percentách. Po dramatickom vývoji v 1. štvrtroku 2001 sa tento podiel znižuje.

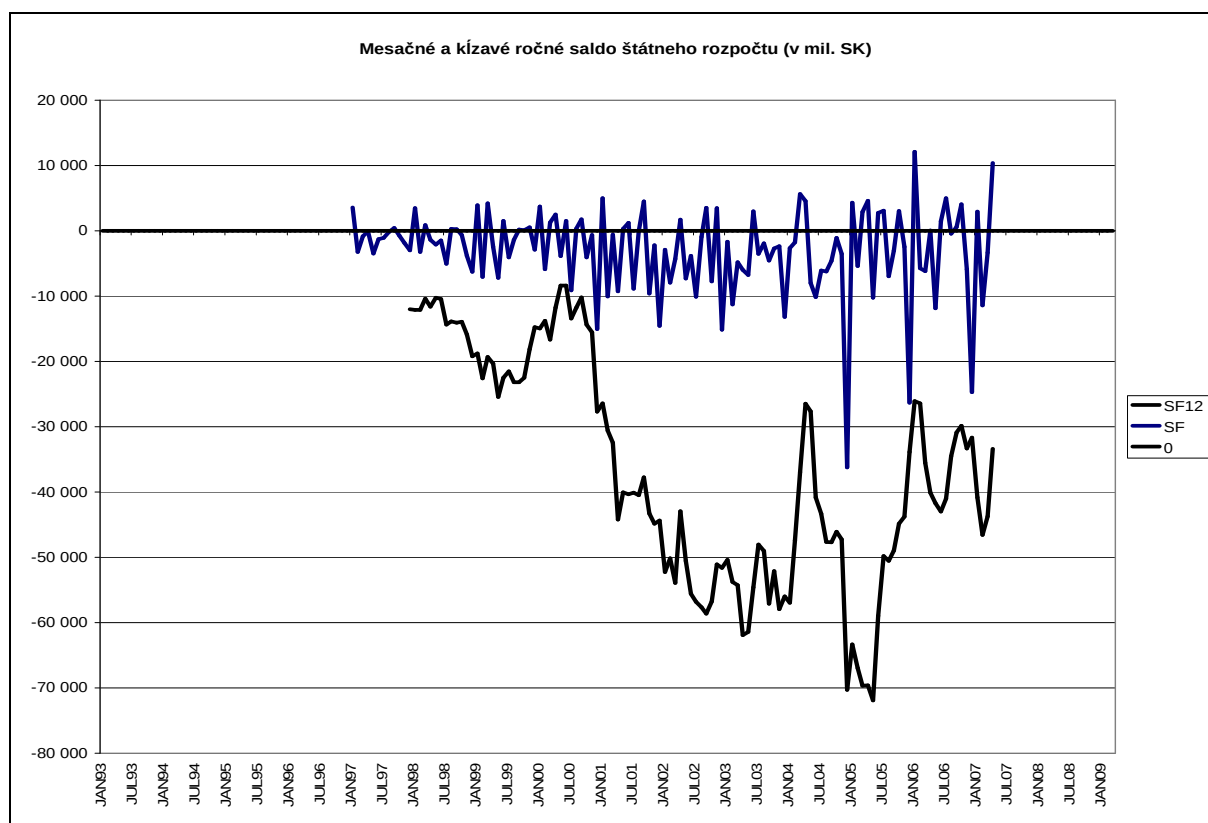
Na obr. 16 je vývoj počtu zamestnaných podľa Výberového zisťovania pracovných síl. Môžeme sledovať rast tohto počtu.



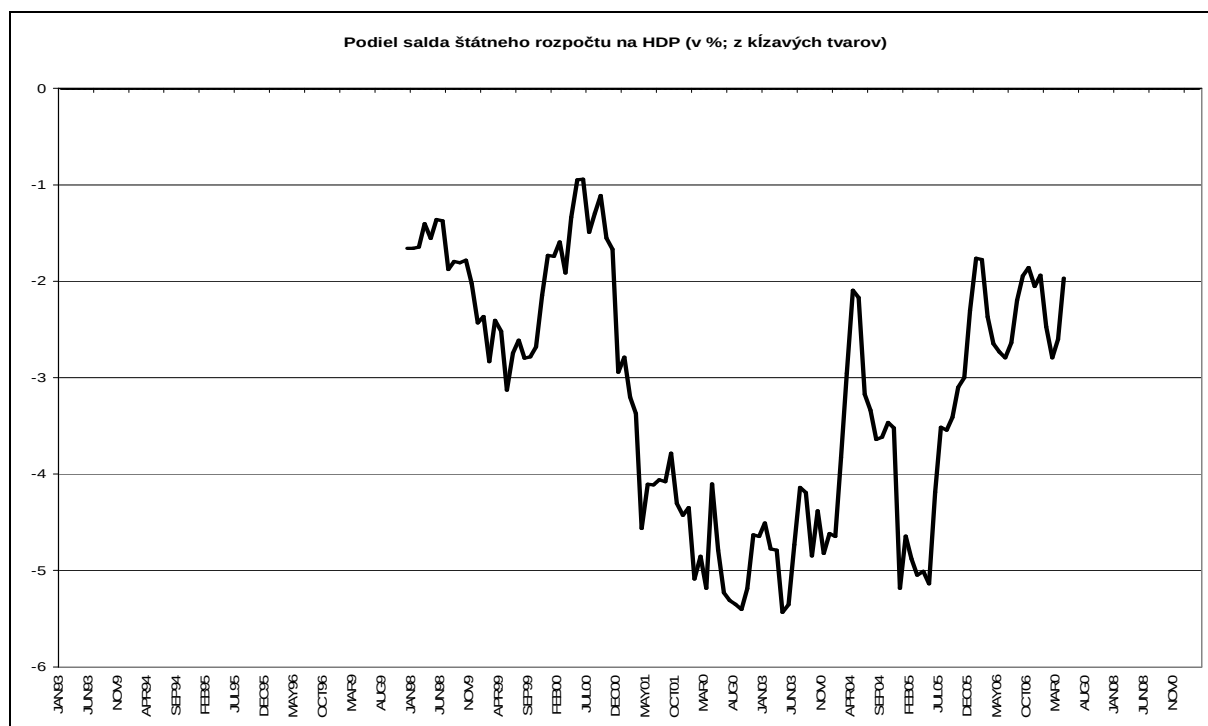
Obr. 1 Medziročný index rastu kĺzavého ročného objemu HDP v s.c. 2000



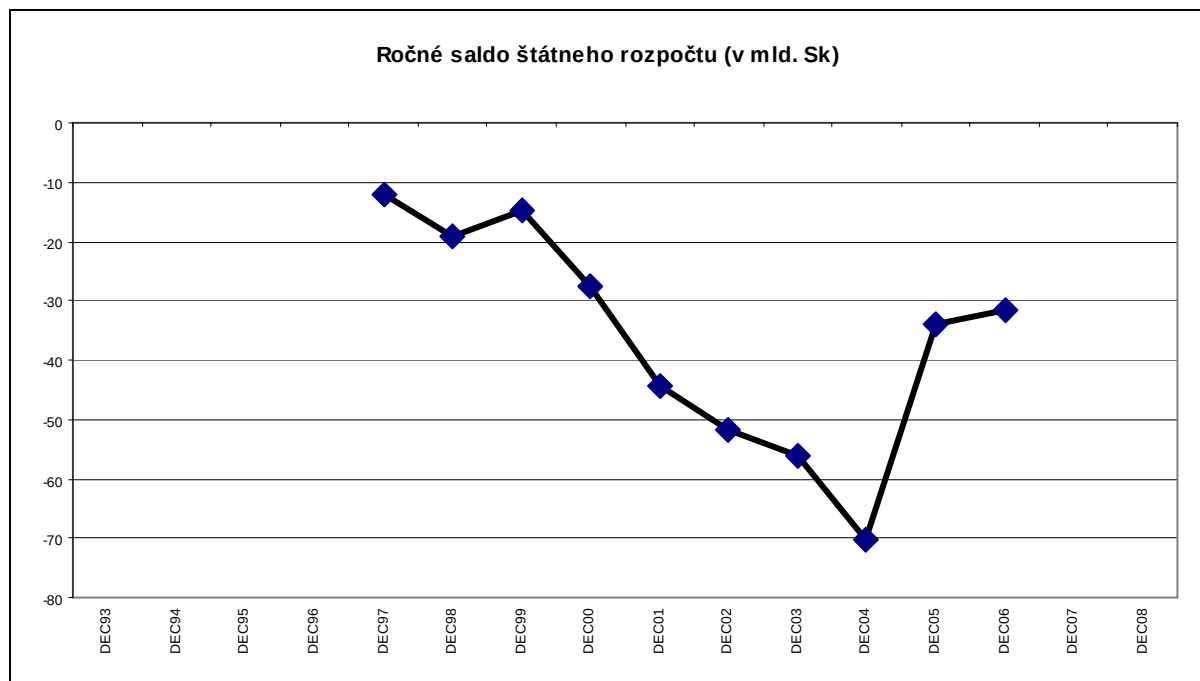
Obr. 2 Medziročný index rastu kĺzavého ročného objemu HDP v bežných cenách



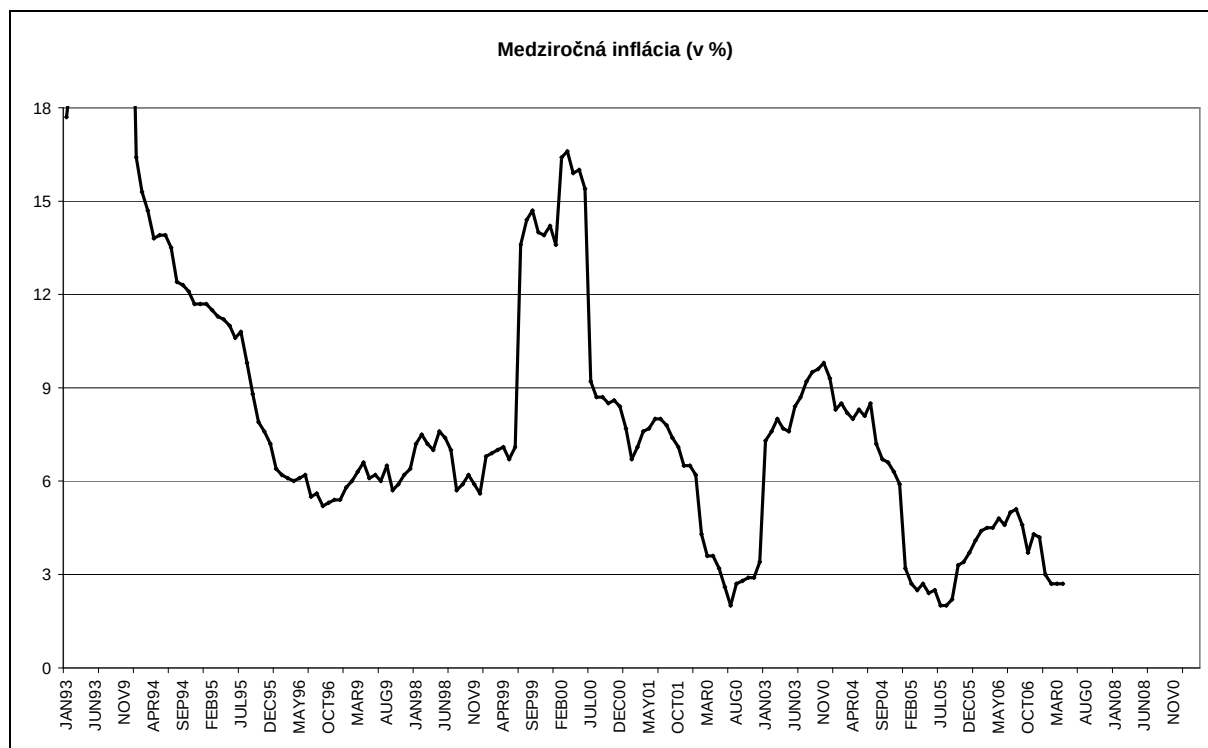
Obr. 3 Mesačné a kľúčové ročné saldo štátneho rozpočtu (v mil. Sk)



Obr. 4 Podiel kľúčového ročného salda štátneho rozpočtu na HDP (v %)



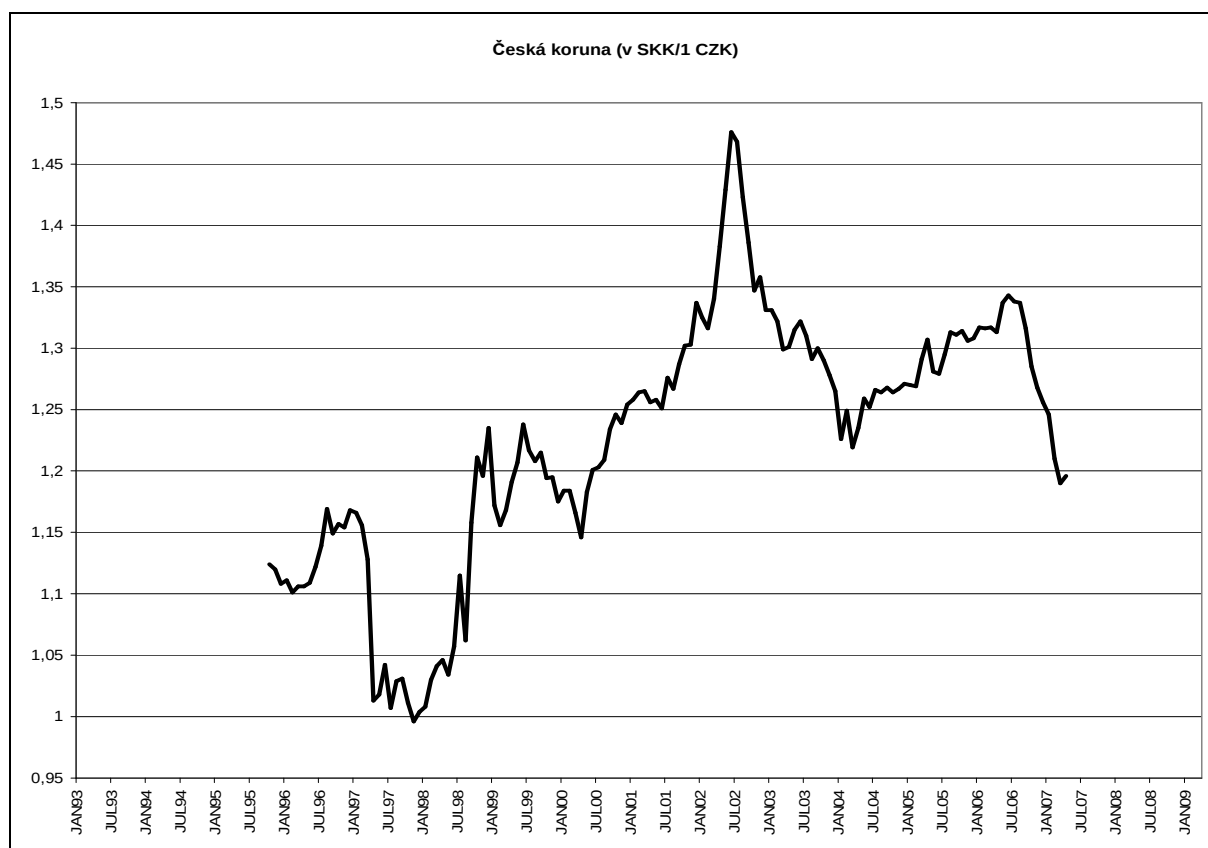
Obr. 5 Ročný objem salda štátneho rozpočtu (v mld. Sk)



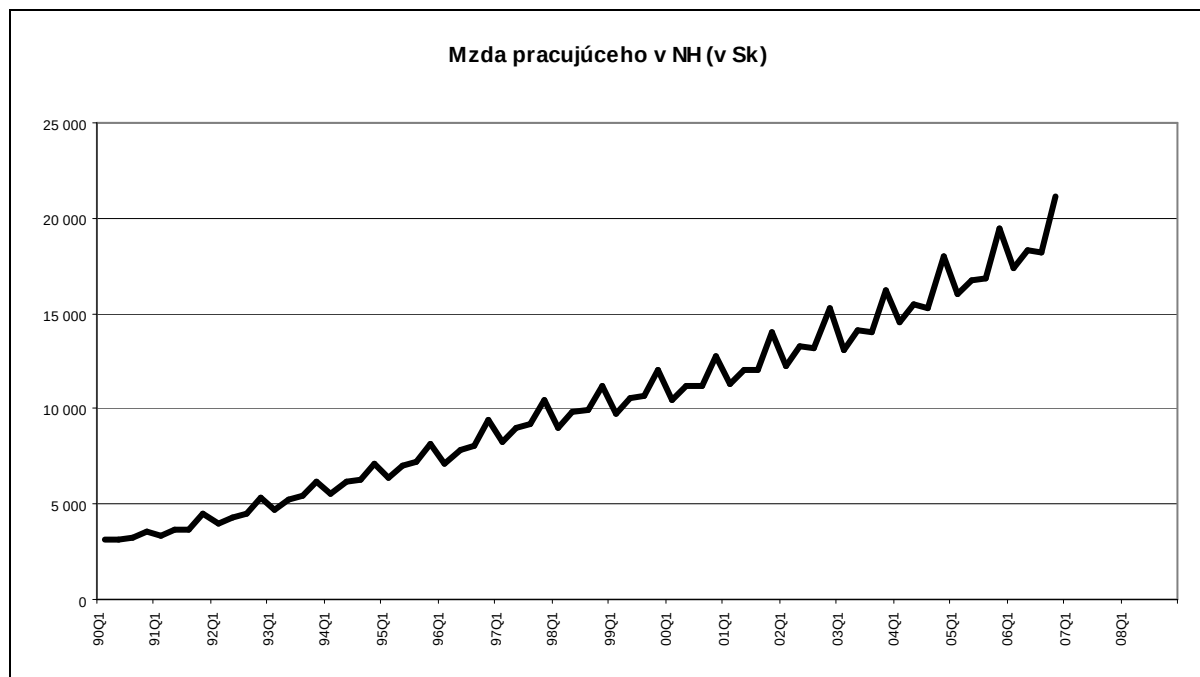
Obr. 6 Medziročná inflácia (v %)



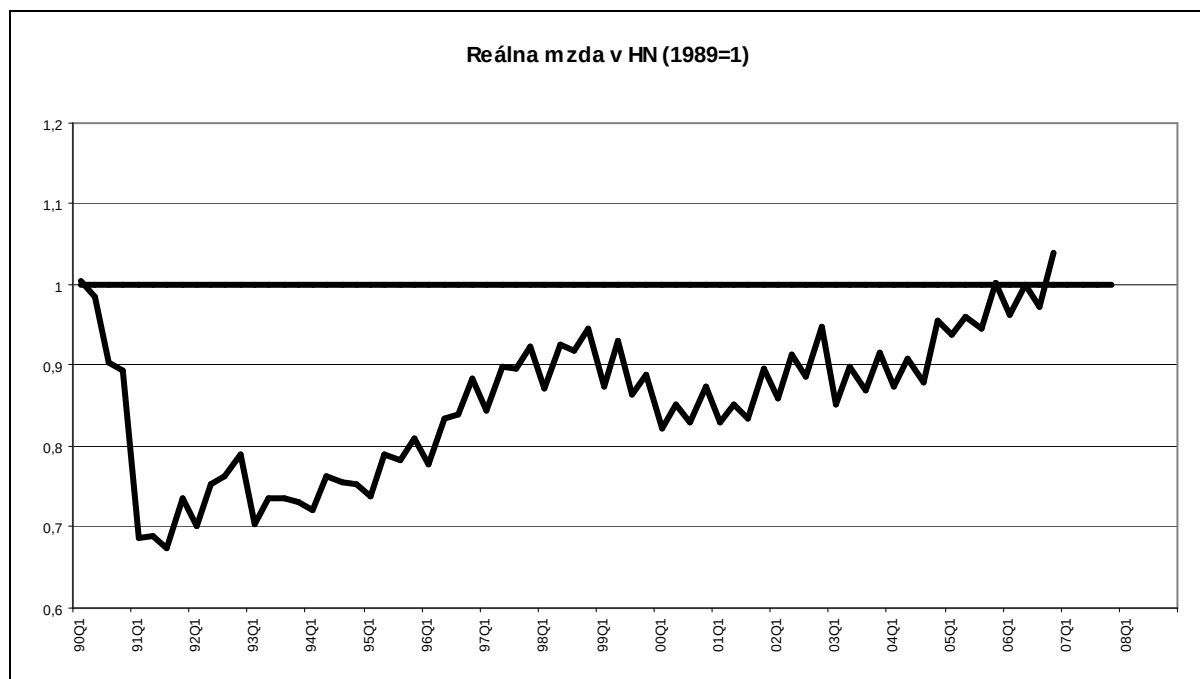
Obr. 7 Kurz eura



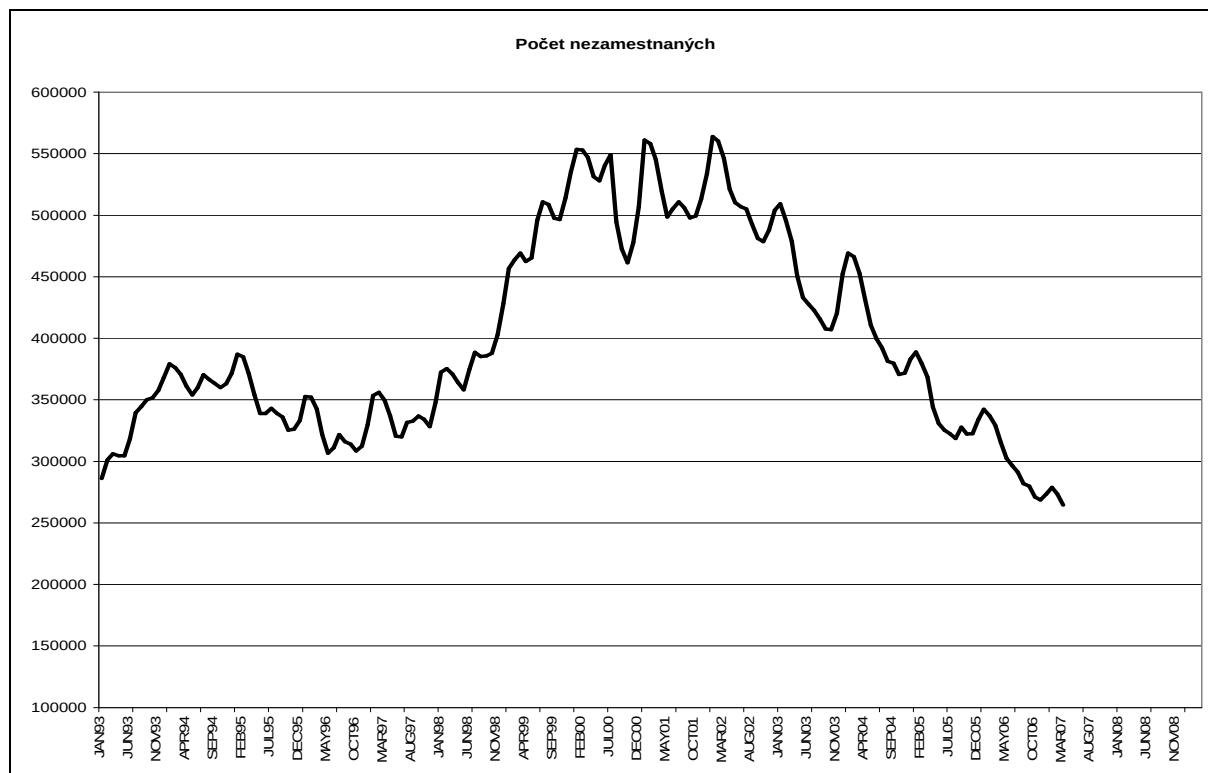
Obr. 8 Kurz českej koruny



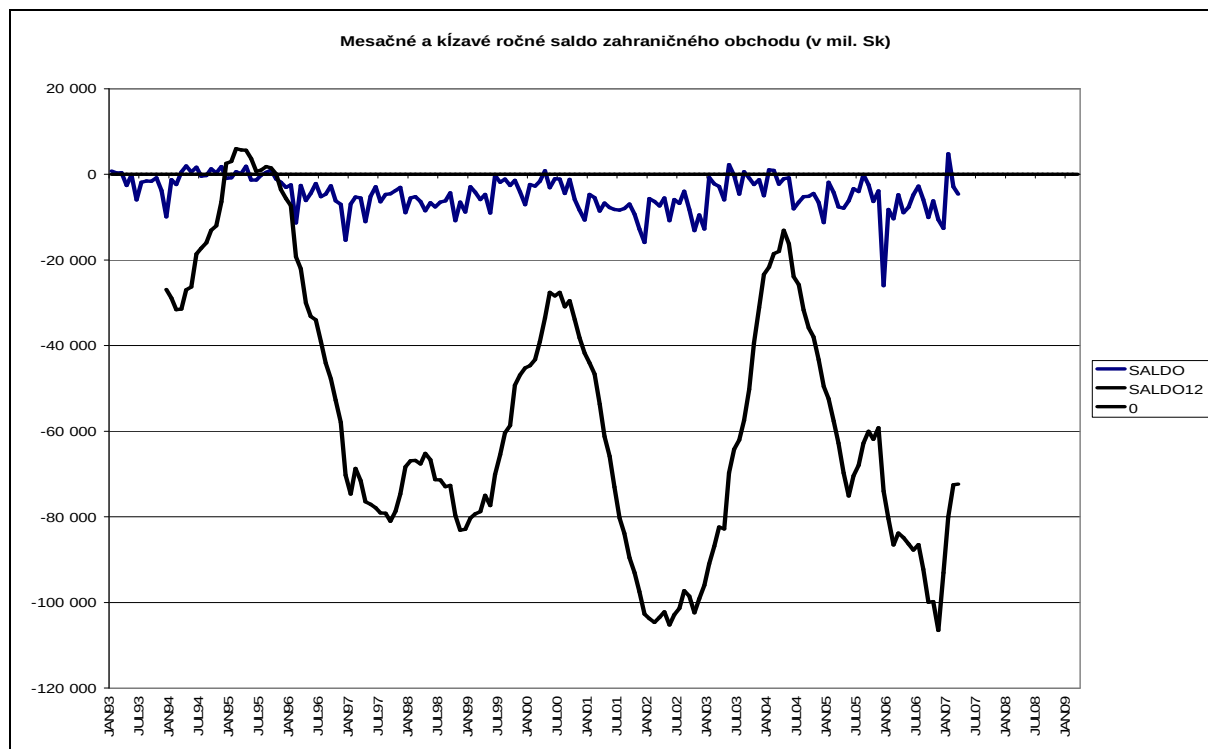
Obr. 9 Priemerná mesačná mzda pracujúceho v národnom hospodárstve



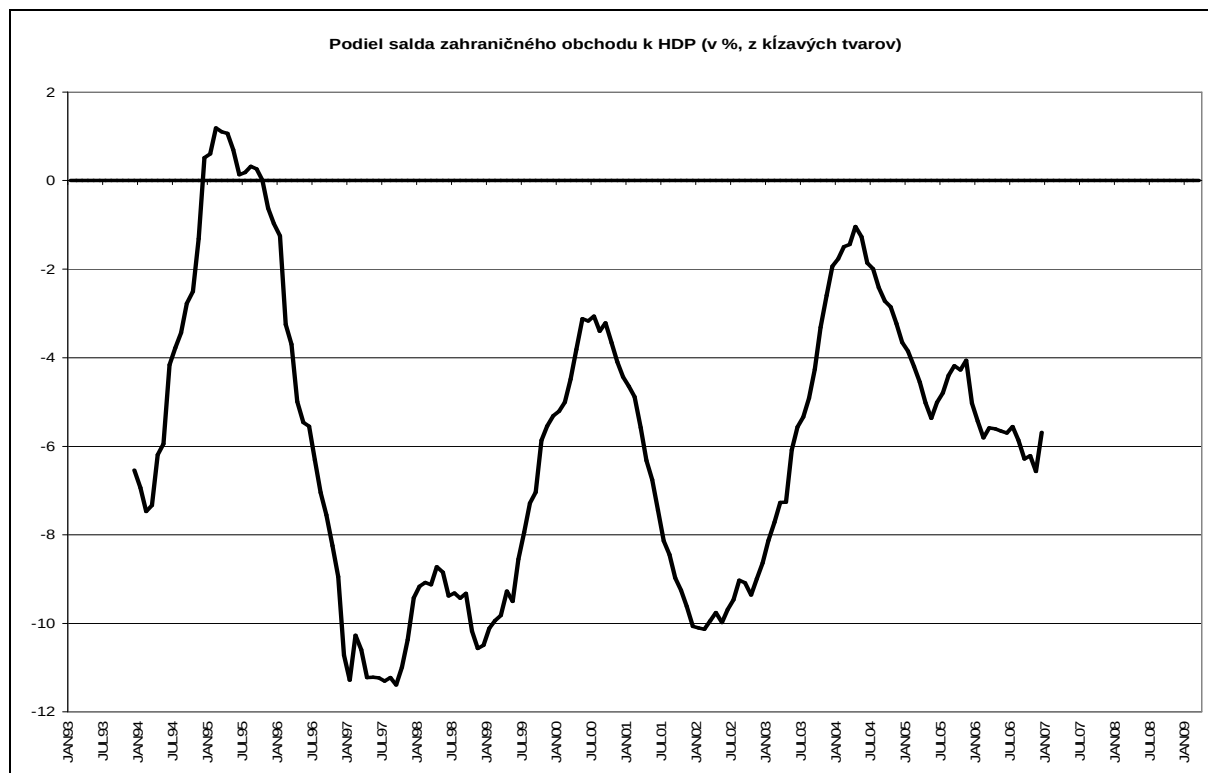
Obr. 10 Bázický index reálnej mzdy pracujúceho v národnom hospodárstve
(štvrtroky 1989 sa rovnajú 1)



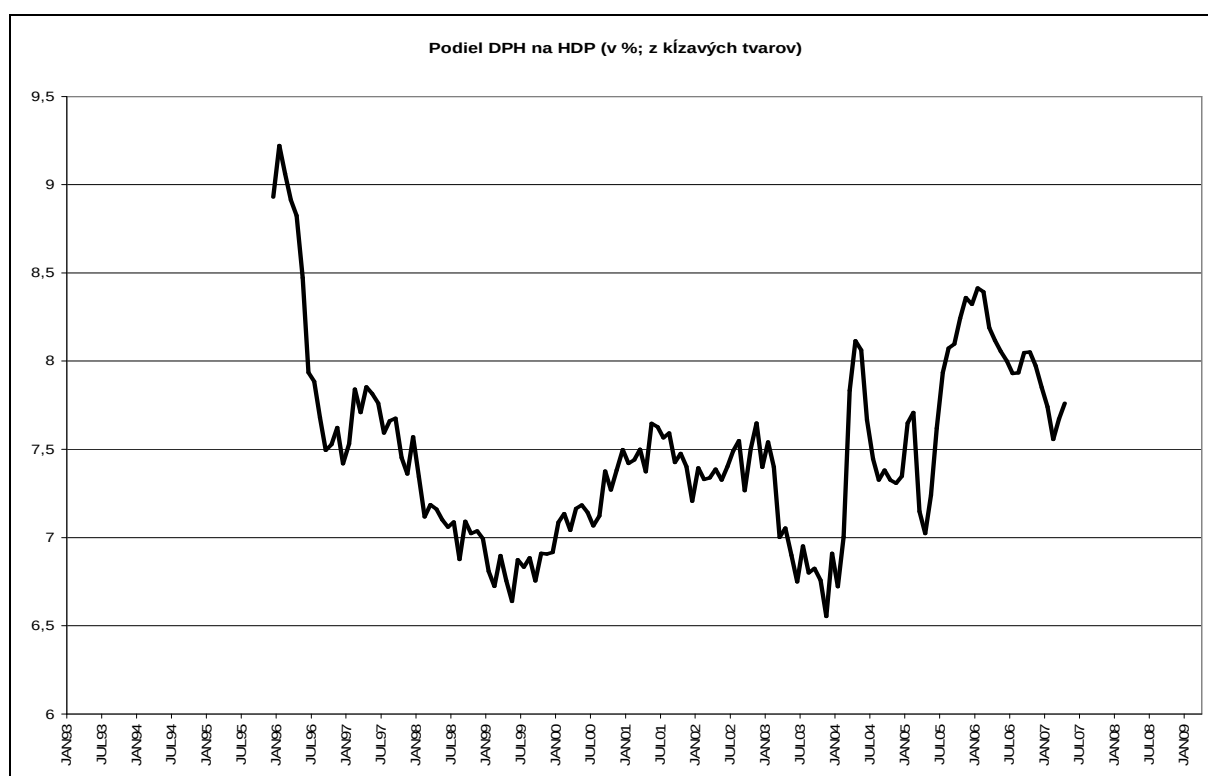
Obr. 11 Celkový počet nezamestnaných



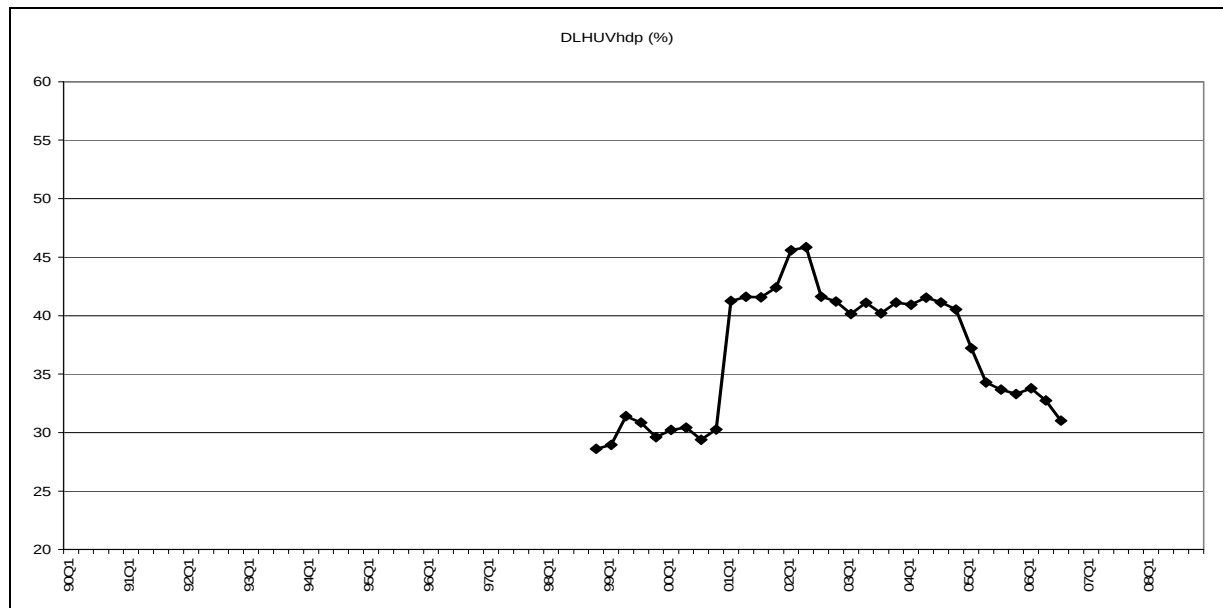
Obr. 12 Mesačné a kĺzavé ročné saldo zahraničného obchodu (v mil. Sk)



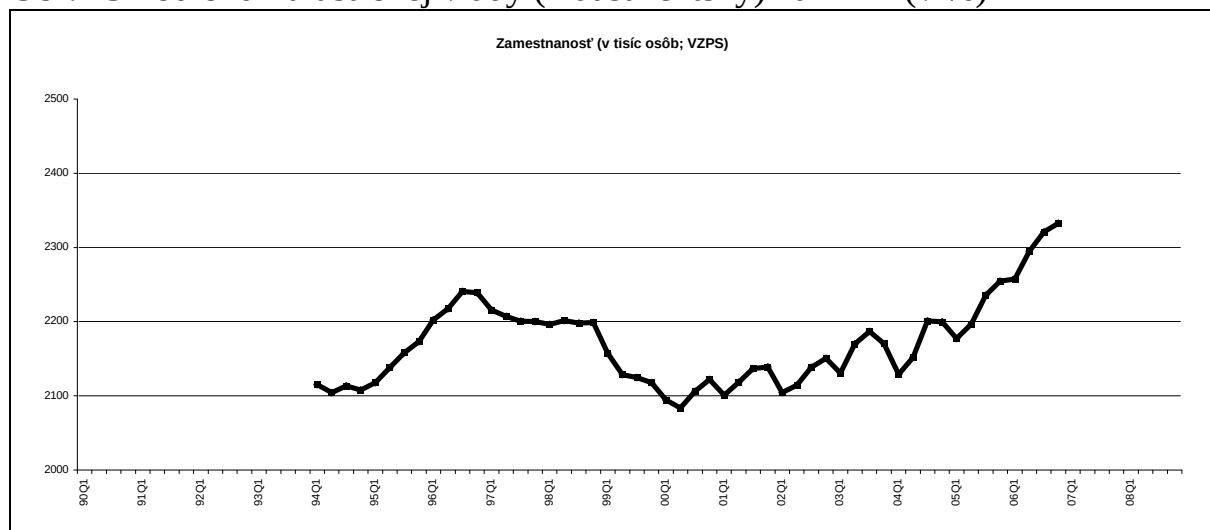
Obr. 13 Podiel kľzavého ročného salda zahraničného obchodu na HDP (v %)



Obr. 14 Podiel kľzavého ročného objemu dane z pridanej hodnoty na HDP (v %)



Obr. 15 Podiel dlhu ústrednej vlády (maastrichtský) na HDP (v %)



Obr. 16 Počet zamestnaných (v tisíc osôb; podľa VZPS)

Zdroj údajov:

EIS Slovensko, Statis, Bratislava

www.statistics.sk; www.nbs.sk; www.finance.gov.sk; www.upsvar.sk

Literatúra:

V. Kvetan, M. Mlýnek, V. Páleník, M. Radvanský: Prognóza vývoja ekonomiky SR na roky 2007 až 2015 s dôrazom na vývoj HDP a zamestnanosti. in: PES 2007, SŠDS Bratislava 2007

Kontakt na autora: chajdiak@statis.biz

Hľadanie skrytej štruktúry pri analýze efektívnosti priemyselných podnikov

Mária Vojtková

Abstract: This article provides a conceptual explanation two techniques data reduction. Principal component analysis is often confused with factor analysis. We will concentrate a discussion of the similarities and the differences between the two techniques. This methods focuses for explanation of relations enterprising data sets on industry of the Slovak Republic. The data are concentrated of economic efficiency.

Key words: Principal component analysis, principal factor analysis, their comparison, the underlying relationships among variables of economic efficiency.

1. Úvod

Metódy zníženia dimenzie, metódy analýzy skrytých vzťahov alebo metódy analýzy nezávislosti je spoločný názov pre skupinu metód, ktoré sa zaoberajú transformáciou súboru napozorovaných premenných do nových hypotetických premenných (hlavných komponentov alebo faktorov), ktoré sú vzájomne nezávislé a sú zoradené podľa veľkosti svojho príspevku k vysvetleniu rozptylu. Pri faktorovej analýze predpokladáme, že merateľné veličiny (vstupné premenné) môžeme vyjadriť ako lineárne funkcie malého počtu skrytých (nemerateľných) – spoločných faktorov a jedného špecifického (chybového) faktora.

Úlohou faktorovej analýzy je nájsť vyjadrenie jednotlivých premenných X_j v tvare:

$$\begin{aligned} X_1 &= a_{11}F_1 + a_{12}F_2 + \dots + a_{1q}F_q + e_1 \\ X_2 &= a_{21}F_1 + a_{22}F_2 + \dots + a_{2q}F_q + e_2 \\ &\vdots \\ X_k &= a_{k1}F_1 + a_{k2}F_2 + \dots + a_{kq}F_q + e_k \end{aligned}$$

kde $F_1, F_2, \dots, F_q$ sú spoločné faktory ($q \leq k$)
 $e_1, e_2, \dots, e_q$ sú špecifické faktory
 a_{jh} sú faktorové saturácie alebo váhy.

Princípy faktorovej analýzy možno zhrnúť do niekoľkých bodov:

- Indikátory (merateľné veličiny, premenné X) sú navzájom korelované, pretože zdieľajú minimálne jeden spoločný znak:
 - ktorý je zodpovedný za koreláciu medzi nimi,
 - nemôže byť priamo zmeraný,
 - pôsobí minimálne na dve premenné X ,
 - nazýva sa spoločný (skrytý) faktor.
- Variabilita indikátorov nevysvetlená skrytým faktorom je spôsobená špecifickými vplyvmi - tzv. špecifickými faktormi resp. náhodnou chybou.

V tomto príspevku sme sa rozhodli aplikovať a porovnať výsledky dvoch najpoužívanejších metód odhadu faktorového modelu a to metódu hlavných faktorov (PFA) a metódu hlavných komponentov (PCA), prípadne poukázať na rozdiely medzi nimi. Metóda hlavných

komponentov sa často uvádza samostatne vzhľadom k odlišnému prístupu pri odhade faktorového modelu, avšak svojou povahou patrí rovnako ako metóda hlavných faktorov k metódam redukcie dát. V úvodnom texte sme sa zamerali predovšetkým na vysvetlenie matematického aparátu metódy PFA, ktorá je akýmsi rozšírením metódy PCA. Samotnú aplikáciu oboch metód sme uskutočnili v SAS Enterprise Guide verzia 4.1.

2. Popis vstupných premenných

Predmetom analýzy sú anonymizované priemyselné podniky, charakterizované ukazovateľmi výstupu z hospodárskeho procesu, poskytnuté pre vedecké účely za rok 2006. Ide o rozsiahly súbor 2367 priemyselných podnikov s počtom zamestnancov viac ako 20. Každý podnik je charakterizovaný pomerne veľkým počtom relatívnych, syntetických ukazovateľov efektívnosti:

Z/VI	- rentabilita vlastného imania,
Z/V	- ziskovosť,
Z/PH	- rentabilita pridanej hodnoty,
PH/ZAM	- produktivita práce z pridanej hodnoty,
V/ZAM	- produktivita práce z výnosov,
PH/ON	- finančná produktivita práce z pridanej hodnoty,
PH/V	- podiel pridanej hodnoty na výnosoch,
CK/ZAM	- vybavenosť zamestnanca celkovým kapitálom,
ON/ZAM	- priemerné osobné náklady na zamestnanca,
CK/VI	- vybavenosť vlastného imania celkovým kapitálom,
ODP/V	- nákladovosť odpisov.

Kde Z je výsledok hospodárenia pred zdanením, VI je vlastné imanie, V sú výnosy, N sú náklady, CK je celkový kapitál (pasíva spolu), PH je pridaná hodnota, ON sú osobné náklady, ODP sú odpisy a ZAM je priemerný počet zamestnancov v prepočítaných osobách.

Všeobecne najvhodnejšie pre faktorovú analýzu sú kvantitatívne premenné v špecifických prípadoch môže ísť aj o poradové alebo umelé (s hodnotami 0-1) premenné. Veľkosť súboru predpokladá, aby počet pozorovaní bol päť až desaťkrát väčší ako počet analyzovaných premenných.

3. Aplikácia metód faktorovej analýzy

a) Posúdenie vhodnosti dát

Keďže faktorová analýza sa najčastejšie používa na definovanie vzťahov medzi premennými, východiskovým zdrojom informácií je obvykle korelačná matica premenných. Pre rovnaké účely je však možné použiť aj Kaiser-Meyer-Olkinovu mieru adekvátnosti výberových dát, ktorú si budeme ďalej označovať ako KMO-index. Ide vlastne o index porovnávajúci veľkosť zistených koeficientov korelácie a veľkosť parciálnych koeficientov korelácie.

KMO - index sa pri všetkých ukazovateľoch efektívnosti pohybuje v intervale od 0,497 po 0,782, t.j. všetky premenné sú vhodné pre použitie faktorovej analýzy. I keď nižšie hodnoty narušujú homogenitu dát, väčšina premenných ako aj dáta celkovo sú hodnotené ako dobré (priemerná hodnota KMO je 0,67647).

b) Výber metódy odhadu faktorov

Faktorová analýza ako viacrozmerná štatistická metóda je výpočtovo náročná, čiže jej použitie je nemysliteľné bez počítačov s vhodným štatistickým softvérom. Pri našej analýze bol použitý Štatistický analytický systém (SAS), takže výber spomenutých metód bol uskutočnený v závislosti od toho, aké možnosti poskytuje tento softvér. SAS ako špičkový štatistický softvér obsahuje celý rad metód extrakcie faktorov. Z nich najpoužívanejšie sú metóda hlavných faktorov a metóda hlavných komponentov. Obe sa používajú na identifikáciu nových dimenzií.

Nedá sa jednoznačne povedať, ktorá z uvedených metód dáva lepšie výsledky. Pri aplikácii týchto metód záleží predovšetkým na skúsenostiach analytika, ktorý uvedené metódy používa. Z toho dôvodu sa tiež tieto metódy považujú skôr za prieskumné (exploratívne) techniky, ktoré môžu byť pomôckou v analytickom uvažovaní.

Metóda hlavných komponentov (PCA) je jedna z najstarších a najviac používaných metód. Cieľom tejto metódy je lineárna transformácia pôvodných premenných na nové, nekorelované premenné – hlavné komponenty, ktoré sú zoradené podľa dôležitosti, t.j. podľa veľkosti svojho príspevku k vysvetleniu celkového rozptylu.

Pri metóde hlavných faktorov (PFA) predpokladáme, že každú vstupnú premennú môžeme vyjadriť ako lineárnu kombináciu nevelkého počtu spoločných skrytých faktorov a jedného špecifického faktora.

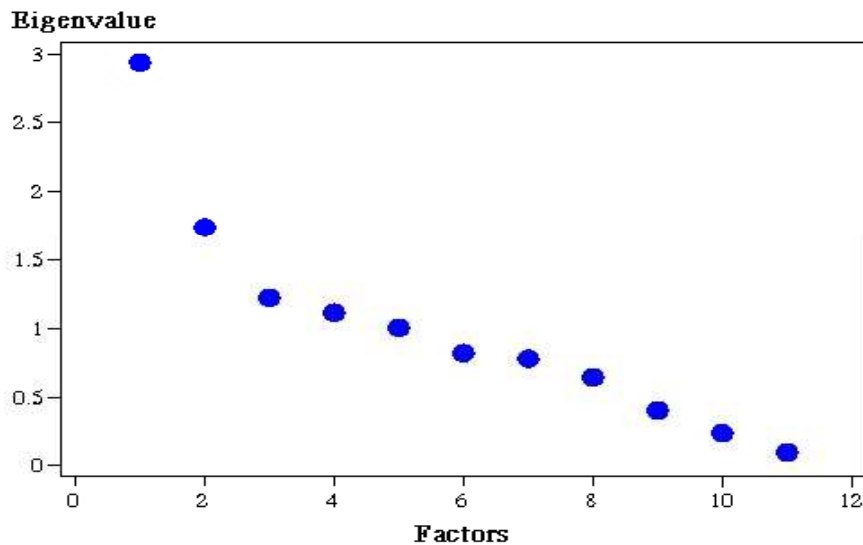
Základný rozdiel medzi týmito metódami je, že PFA vysvetľuje koreláciu medzi premennými a PCA vysvetľuje variabilitu premenných. Pri aplikácii metódy PCA používame korelačnú maticu, ktorá má na diagonále jednotky, čiže celkový rozptyl. Pri metóde PFA sa na diagonálu korelačnej matice dosadzujú komunality, t.j. časť rozptylu, ktorú je možné vysvetliť pôsobením spoločných faktorov.

V metóde PCA sú premenné označené ako formujúce (tvoriace) indikátory hlavných komponentov. Na druhej strane v metóde PFA premenné odrážajú prítomnosť skrytej štruktúry faktorov. Inými slovami premenné sú označené ako odrážajúce indikátory.

V tomto článku sme aplikovali najprv metódu hlavných komponentov, ktorú sme následne porovnali s výsledkami iteračnej metódy hlavných faktorov.

c) Určenie počtu významných faktorov

Faktorová analýza umožňuje odhadnúť maximálne toľko faktorov, koľko je vstupných, merateľných premenných. Cieľom však je vysvetliť maximum korelácie (rozptylu) medzi premennými pomocou minimálneho počtu faktorov. Vzniká problém, koľko faktorov z odhadnutého počtu stačí na vysvetlenie korelácie medzi premennými. Analytik sa musí rozhodnúť, akú stratu informácií je ochotný akceptovať. V exaktných vedách sa vyžaduje, aby 90 - 95% variability pôvodných premenných bolo vysvetlené pôsobením spoločných faktorov, v spoločenských vedách aspoň 60%. Podporne možno použiť aj tzv. scree plot, ktorý zobrazuje počet faktorov verzus ich vlastné čísla. V grafe sa snažíme nájsť zlom a akceptujeme faktory, ktoré sa nachádzajú pred týmto zlomom.



Obrázok 1: Graf vlastných čísel (scree plot)

V našej analýze sme tiež postupovali na základe tabuľky vlastných čísel (tabuľka 1) a grafu vlastných čísel. Vlastné číslo charakterizuje rozptyl každého faktora, pričom celkový rozptyl vzhľadom k tomu, že pracujeme so štandardizovanými premennými je rovný 11. Najviac nás bude zaujímať posledný stĺpec v tabuľke 1, ktorý vyjadruje kumulatívny podiel variability vysvetlený daným počtom faktorov. Ďalším kritériom pre výber štatisticky významných faktorov je porovnanie vlastných čísel s priemerným vlastným číslom, čiže 1. Celkovo teda môžeme uvažovať s 5 faktormi, ktoré vysvetľujú 72,89% celkovej variability. Ďalší zlom možno vidieť pri 7 faktoroch, ktoré vysvetľujú až 87,43 % variability, avšak ich počet vzhľadom k počtu vstupných premenných je už pomerne vysoký. Za postačujúci budeme teda považovať počet faktorov päť. Toto rozhodnutie sme si overili aj pomocou grafu vlastných čísel (obrázok1).

Tabuľka 1: Vlastné čísla korelačnej matice (metóda PCA)

Eigenvalues of the Correlation Matrix: Total = 11 Average = 1				
	Eigenvalue	Difference	Proportion	Cumulative
1	2.93667802	1.1968023	0.2670	0.2670
2	1.73987565	0.5171125	0.1582	0.4251
3	1.22276309	0.1086231	0.1112	0.5363
4	1.11413996	0.1095158	0.1013	0.6376
5	1.00462410	0.1845504	0.0913	0.7289
6	0.82007369	0.0406736	0.0746	0.8035
7	0.77940003	0.1353788	0.0709	0.8743
8	0.64402123	0.2410513	0.0585	0.9329
9	0.40296993	0.1647541	0.0366	0.9695
10	0.23821577	0.1409772	0.0217	0.9912
11	0.09723853		0.0088	1.0000

d) Rotácia faktorov

Faktorová analýza umožňuje zoskupenie premenných do nových dimenzií, faktorov. Pôvodné premenné sú do nových faktorov spájané pomocou numerických váh. Tieto váhy sa nazývajú faktorové (komponentné) saturácie. Sú kľúčom k pochopeniu a interpretácii jednotlivých faktorov.

Ak nemožno nerotované riešenie jednoznačne považovať za vyhovujúce, je vhodné použiť metódy rotácie faktorov na jeho zjednodušenie. Spoločným cieľom všetkých metód rotácie je dosiahnuť lepšiu interpretáciu riešenia pomocou zjednodušenia riadkov a stĺpcov matice faktorovej štruktúry. Pri rotácii dochádza k otáčaniu osí jednotlivých faktorov, pričom sa nemenia odhady konečných komunalít jednotlivých premenných získaných z nerotovaného riešenia ako aj podiel celkovej variability vysvetlený spoločnými faktormi. Rotáciou sa iba prerozdeľuje vysvetlená variabilita medzi jednotlivými faktormi. Existujú dva základné typy rotácie:

- ortogonálna – je najjednoduchším typom rotácie, ktorá zachováva nekorelovanosť premenných. Je vhodná pre redukciu veľkého počtu premenných.
- kosouhlá – jej výsledkom sú korelované premenné, avšak kosouhlé riešenia sú často flexibilnejšie a realistickejšie.

Za účelom získania interpretovateľnejšieho riešenia bola v našej analýze použitá ortogonálna rotácia VARIMAX. Ide o rotáciu zameranú na zjednodušenie stĺpcov matice faktorovej štruktúry tým, že minimalizuje počet premenných vysoko korelovaných s jednotlivými faktormi. Dosiahne sa tým faktorový model, v ktorom každý faktor reprezentuje rozdielnú skutočnosť, čiže jednoduchšiu interpretáciu. Pri väčšom počte premenných zvolený typ ortogonálnej rotácie má zanedbateľný vplyv na celkový výsledok.

e) Interpretácia faktorov

Interpretácia odhadnutých faktorov a celého modelu vychádza z faktorových (komponentných) saturácií jednotlivých faktorov po rotácii a zo znalostí samotného analytika o ich vzájomných vzťahoch. Vysoká hodnota saturácie znamená, že faktor a daná premenná majú veľa spoločného. Na jednej strane daná premenná dobre identifikuje skrytý faktor a na strane faktor významne ovplyvňuje variabilitu danej premennej.

Tabuľka 2: Rotovaná faktorová matica (metóda PCA)

Rotated Factor Pattern					
	Factor1	Factor2	Factor3	Factor4	Factor5
V_ZAM	0.87189	0.03845	-0.00240	-0.04467	0.03084
CK_ZAM	0.77709	0.14151	-0.00493	0.13513	0.02639
ON_ZAM	0.73052	-0.03849	0.00892	-0.15199	-0.02855
PH_ON	0.22008	0.92376	-0.00026	-0.05685	-0.00057
PH_V	-0.22428	0.88395	-0.00083	-0.08893	0.00517
PH_ZAM	0.59825	0.73631	0.00006	-0.01878	-0.00870
CK_VI	-0.00494	0.00984	0.78158	0.01309	0.01244
Z_VI	0.00613	-0.01013	0.78045	-0.02194	-0.01239
ODP_V	0.04520	0.04354	0.00925	0.85864	-0.15345
Z_V	0.12931	0.22980	0.02366	-0.59607	-0.26212
Z_PH	0.03669	0.01775	0.00268	0.01341	0.96144

Výsledky rotácie modelu faktorovej analýzy sú prehľadne zobrazené v tabuľke 2, pričom váhy jednotlivých faktorov sú usporiadané podľa veľkosti. Skupina premenných s veľmi významnými faktorovými saturáciami predstavuje najlepšie indikátory faktora vo vytvorenom modeli. Znamená to, že ich variabilita je najviac ovplyvňovaná zmenami daného faktora.

Ako prvú sme aplikovali metódu hlavných komponentov, pričom podľa faktorových saturácií¹ jednotlivé faktory možno interpretovať ako nasledovné skryté dimenzie:

1. faktor – produktivita práce a vybavenosť zamestnanca,
2. faktor - produktivita práce z pridanej hodnoty a jej štruktúra,
3. faktor – vybavenosť a rentabilita vlastného imania,
4. faktor – nákladovosť odpisov verzus ziskovosť,
5. faktor – rentabilita pridanej hodnoty.

Vytieňované faktorové saturácie poukazujú na koreláciu daného faktora a danej premennej. Väčšina zvýraznených faktorových saturácií je kladná, iba pri štvrtom faktore sa môžeme stretnúť s kladnou aj zápornou hodnotou váh. Tento faktor teda naznačuje dve rozdielne orientácie efektívnosti a to kladnú (zvyšovanie nákladovosti) alebo zápornú (znižovanie ziskovosti). O niečo vyššia váha je pri kladnej orientácii, avšak podniky uprednostňujú buď jednu alebo druhú, ale nie obidve naraz.

Podiel celkovej vysvetlenej variability je približne 73 %, čiže vybrané spoločné faktory dostatočne reprodukovujú skryté vzťahy medzi pôvodnými premennými. Zvyšok variability je možné prisúdiť vplyvu špecifických faktorov.

Tabuľka 3: Rotovaná faktorová matica (metóda PFA)

Rotated Factor Pattern					
	Factor1	Factor2	Factor3	Factor4	Factor5
PH_ON	0.93490	0.19918	0.05616	0.00053	-0.01352
PH_V	0.77250	-0.19985	0.10298	-0.00185	-0.09764
PH_ZAM	0.75001	0.59283	0.00340	0.00249	0.09306
V_ZAM	0.06008	0.88761	-0.15738	0.00095	-0.10835
CK_ZAM	0.14060	0.64990	-0.01364	-0.00523	0.28123
ON_ZAM	0.03038	0.56297	0.01757	0.00826	-0.03443
Z_V	0.14074	0.19639	0.81243	0.00770	-0.15170
Z_PH	0.00226	0.02922	-0.07713	-0.00291	0.01781
Z_VI	-0.00480	0.00562	0.01283	0.48036	-0.00868
CK_VI	0.00443	-0.00227	0.00467	0.45887	0.00148
ODP_V	-0.02383	-0.01546	-0.09226	-0.00318	0.36370

Pre porovnanie bola uskutočnená aj metóda hlavných faktorov s rotáciou Varimax, pričom počet významných faktorov bol určený v súlade s riešením získaným pomocou metódy hlavných komponentov. Pri odhade počiatočných komunalít sme v našom prípade použili štvorec viacnásobného koeficienta korelácie podľa jednotlivých premenných. Na získanie prijateľného riešenia bolo nutné použiť 222 iterácií.

Výsledné riešenie pre prvé dva faktory sa zhoduje s riešením pomocou metódy PCA i keď v opačnom poradí. Tretí faktor súvisí už iba s ukazovateľom ziskovosti. Všetky faktorové saturácie pre ostatné dva faktory sú pomerne nízke, čo poukazuje na ich nevýznamnosť. Tento

¹ Pri metóde PCA sa nové premenné označujú aj ako hlavné komponenty a ich váhy ako komponentné saturácie.

problém je možné riešiť znížením počtu významných faktorov, čo má za následok aj zníženie celkového množstva vysvetlenej variability premenných.

V záujme zachovať vo faktorovom modeli, čo najväčšie množstva informácií o pôvodnom súbore sa javí toto riešenie ako neprijateľné. Pre porovnanie sme aplikovali metódu PFA aj použitím iných inicializačných odhadov komunalít avšak získané výsledky boli podobné, takže ich vzhľadom k rozsahu nebudeme uvádzať.

Tabuľka 4: Odhad rozptylu vysvetlený 5-faktorovým modelom metódou PCA

Final Commuality Estimates: Total = 8.018081						
PH_ZAM	V_ZAM	PH_ON	PH_V	Z_VI	Z_V	Z_PH
0.90048174	0.76462089	0.90500633	0.83960976	0.60988052	0.49410271	0.92621455

CK_ZAM	ON_ZAM	CK_VI	ODP_V
0.64287650	0.55913682	0.61130779	0.76484322

Tabuľka 5: Odhad rozptylu vysvetlený 5-faktorovým modelom metódou PFA

Final Commuality Estimates: Total = 5.497259						
PH_ZAM	V_ZAM	PH_ON	PH_V	Z_VI	Z_V	Z_PH
0.92263755	0.82797553	0.91705607	0.65684508	0.23104266	0.74148940	0.00713438

CK_ZAM	ON_ZAM	CK_VI	ODP_V
0.52143764	0.31942269	0.21061182	0.14160664

Porovnaním uvedených výsledkov sa prijateľnejšie a interpretovateľnejšie javí riešenie získané metódou hlavných komponentov. Tento záver potvrdzuje aj porovnanie tabuliek 4 a 5. Pri jednotlivých premenných je vo väčšine prípadov podiel variability vysvetlený uvedenými 5 faktormi vyšší pri metóde PCA ako pri PFA. Napríklad pre rentabilitu pridanej hodnoty pri metóde PCA tvorí podiel variability vysvetlený 5-faktorovým modelom až 92,6%, zatiaľ čo pri metóde PFA tvorí iba 0,7 %. Zvyšnú časť variability považujeme za nevysvetlenú, pričom jej veľkosť je možné prisúdiť vplyvu špecifických faktorov a chybovosti. Podobne pomerne nízke hodnoty variability vysvetlenej 5-faktorovým modelom možno pri metóde PFA pozorovať aj pri ukazovateľoch nákladovosť odpisov, vybavenosť a rentabilita vlastného imania.

4. Záver

Faktorová analýza umožňuje nájsť logickú kombináciu premenných a tak lepšie pochopiť vnútorné väzby medzi premennými. Konštrukciou faktorovej matice a interpretáciou významných faktorov je možné analýzu ukončiť. Ak je našim cieľom použiť identifikované, vhodné faktory pre následnú aplikáciu iných štatistických techník, analýza pokračuje výpočtom faktorových skóre. Ide o kombinovanú, zmiešanú mieru každého faktora, vyčíslenú pre každý objekt (podnik). Výhodou použitia tejto miery je, že zastupuje všetky premenné, ktoré boli s faktorom skorelované, jej interpretácia je však pomerne komplikovaná.

V tomto príspevku išlo predovšetkým o pochopenie vzájomnej, skrytej štruktúry medzi vstupnými ukazovateľmi efektívnosti a následne o ich redukciu. Pri analýze boli použité dve metódy, pričom v danom prípade sa javí ako vhodnejšia metóda hlavných komponentov. Na záver treba pripomenúť, že uvedená analýza patrí medzi prieskumné techniky, ktoré môžu

ďalej viesť napríklad k sledovaniu podobnosti priemyselných podnikov, alebo k modelovaniu závislosti nových, hypotetických premenných.

5. Literatúra

- CHAJDIAK, J.: *Ekonomická analýza stavu a vývoja firmy*. Bratislava: Statis, 2004.
- KHATTREE, R. – NAIK, N. D.: *Multivariate Data Reduction and Discrimination with SAS® Software*. First edition, Cary, NC: SAS Institute Inc., 2000.
- MELOUN, M. – MITICKÝ, J.: *Statistická analýza experimentálnych dat*. Praha: ACADEMIA, 2004.
- SHARMA, S.: *Applied multivariate techniques*. New York: John Wiley & Sons, 1996.
- STANKOVIČOVÁ, I.: *Ako postupovať pri faktorovej analýze v SAS Enterprise Guide*. Forum Statisticum Slovacum. - ISSN 1336-7420. - Roč. 2, č. 5 (2006), s. 135-139.

Adresa autora:

Mária Vojtková, Ing. PhD.
Katedra štatistiky, FHI, Ekonomická univerzita
Dolnozemska 1/b
852 35 Bratislava
E-mail: vojtkova@euba.sk

Príspevok bol spracovaný v rámci riešenia grantovej úlohy VEGA 1/2631/05 „Analýza možností aplikácie viacrozmerných štatistických metód na skúmanie ekonomických výsledkov na príklade priemyslu SR prípadne iných oblastí ekonomiky“.

Vybrané metody určovania váh kritérií pri analyzovaní postavenia podniku na trhu

Vladimír Úradníček

Abstract: The contribution deals with partial problem of firm market position analysis. We apply selected criterion weights design methods to estimate the financial criterion weights used to measure the position of selected firm in relevant market. Applied on real-life data we illustrate the usefulness of pair-wise criterion comparison matrix method, approximate and exact Saaty's criterion weights estimation procedure.

Key words: analýza postavenia podniku na trhu, váha kritérií, metódy párového porovnávania, komparačná matica, Saatyho postup.

1. Úvod

Na analyzovanie postavenia podniku na trhu môžeme variantne používať rôzne typy metód. V závislosti od podstaty použitej metódy sa zvyknú klasifikovať:

- a) *Kvantitatívne metódy,*
- b) *Komparatívno-analytické metódy,*
- c) *Grafické metódy – napr. pavučinové (radarové) grafy, Chernoffove tváre a p.,*
- d) *Kombinácia kvantitatívnych, komparatívno-analytických a grafických metód.*

V procese realizácie analýzy musíme riešiť, okrem iného, tiež problémy, ktoré súvisia s:

- koncipovaním súboru komparovaných podnikov,
- výberom a prípravou relevantných dát a indikátorov (ukazovateľov – kritérií),
- určením charakteru ukazovateľa vo väzbe na hlavný cieľ analýzy a v neposlednom rade aj
- so stanovením váh jednotlivých ukazovateľov – kritérií.

2. Teoretické východiská a aplikácia metód

Pri charakterizovaní postavenia sledovaného podniku na trhu môžeme identifikovať nasledovné etapy analýzy:

1. Stanovenie cieľa, obsahu a rozsahu medzipodnikového porovnávania.
2. Koncipovanie výberového súboru porovnávaných podnikov.
3. Voľba príslušnej (analytickej) metódy, resp. metód.
4. Výber vhodných indikátorov pre proces porovnávania.
5. Príprava dát a jednotlivých indikátorov.
6. Použitie príslušnej (analytickej) metódy, resp. metód.
7. Zhrnutie a interpretácia získaných výsledkov.
8. Variantné návrhy opatrení, vyplývajúcich z dosiahnutých záverov – s uvedením príslušného stupňa stochasticity
9. Vyhodnotenie dôsledkov realizovania, resp. nerealizovania zvolených opatrení – „spätná väzba“

Kvantifikáciu váh vybraných indikátorov (najčastejšie ide o ukazovatele) môžeme uskutočniť viacerými metódami, napríklad pomocou delfskej metódy, metódou SPAN, metódami párovej komparácie – komparačnou maticou, Saatyho metódou a pod.

V príspevku uvidíme praktickú ukážku možnosti stanovenia váh ukazovateľov pomocou metódy SPAN a vybranými typmi metód párovej komparácie.

Metóda SPAN (Successive Proportional Additive Numeration)¹ je počítačom podporovaný proces expertného posúdenia jednotlivých variantov projektu. Modifikovane môžeme túto metódu použiť aj pre stanovenie váh hodnotiacich kritérií, ktorými sú v našom prípade finančné ukazovatele. Princíp metódy SPAN spočíva v tom, že každej expertnej skupine je pridelený rovnaký počet bodov (napríklad 100). Každá expertná skupina rozdelí svoje body na dve časti. Prvá časť je určená ostatným skupinám a druhá časť je určená pre ohodnotenie významnosti („váhy“) danej množiny finančných ukazovateľov. Pomer medzi rozdelenými bodmi zostáva počas celého hodnotiaceho procesu konštantný. V ďalšom kroku každá expertná skupina rozdelí svoje body pridelené skupinám expertov medzi jednotlivé expertné skupiny a body pridelené množine ukazovateľov medzi jednotlivé ukazovatele. Určené pomery rozdelených bodov sú opäť pre celý proces nemenné. V uvedenom pridelovaní bodov sa pokračuje v ďalších kolách až do okamihu, kedy je súčet bodov pridelených ukazovateľov prakticky rovný celkovému počiatočnému počtu všetkých bodov. Následne sa získané hodnoty „zoštandardizujú“. Príkladom je stanovenie váh ukazovateľov: ROE (rentabilita vlastného kapitálu), ROS (rentabilita tržieb), T/A (obrátka majetku v tržbách) a DSKZ (doba splácania krátkodobých záväzkov). Hodnotenie realizovalo šesť rôznych expertných skupín. Postup riešenie a výsledok je uvedený v tabuľkách 1 a 2.

Tabuľka 1 Metóda SPAN

Expert	Pridelené body												
	ROE	ROS	T/A	DSKZ	ukaz. spolu	E1	E2	E3	E4	E5	E6	experti spolu	body spolu
E1	15	12	9	4	40	×	10	12	12	12	14	60	100
E2	19	16	10	5	50	10	×	10	8	11	11	50	100
E3	35	13	16	6	70	6	6	×	6	6	6	30	100
E4	22	15	12	11	60	8	8	8	×	8	8	40	100
E5	18	14	12	6	50	9	8	12	8	×	13	50	100
E6	20	12	11	7	50	9	7	11	14	9	×	50	100
Spolu	129	82	70	39	320	42	39	53	48	46	52	280	600

Prameň: Vlastné spracovanie

Tabuľka 2 Výsledné stanovenie váh kritérií metódou SPAN

Expert	ROE	ROS	T/A	DSKZ	Spolu
E1	26,5	21,2	15,9	7,1	70,8
E2	32,6	27,5	17,2	8,6	85,9
E3	68,4	25,4	31,3	11,7	136,9
E4	41,2	28,1	22,5	20,6	112,4
E5	33,0	25,7	22,0	11,0	91,6
E6	38,6	23,2	21,2	13,5	96,6
Spolu	240,5	151,1	130,1	72,5	594,2
v_j	0,41	0,25	0,22	0,12	1,00

Prameň: Vlastné spracovanie

¹ Bližšie viď: Fiala, P.: Skupinové rozhodování. Praha, VŠE 1997, 193 s.

Medzi najčastejšie používané metódy párového porovnávania patria:

- Matica párových porovnaní,
- Saatyho metóda.

Matica párových porovnaní (komparačná matica) je štvorcová matica typu $m \times m$. Pre každé kritérium sa zisťuje počet jeho preferencií vzhľadom ku všetkým ostatným kritériám súboru.

Ak preferujeme kritérium uvedené v riadku pred kritériom uvedeným v stĺpci napíšeme do príslušného políčka 1, v opačnom prípade 0.

Pre každé kritérium sa stanoví počet jeho preferencií f_j , ktorý sa rovná súčtu jednotiek v riadku uvažovaného kritéria zväčšeného o počet núl v stĺpci tohto kritéria. Na základe počtu preferencií jednotlivých kritérií sa ich normované váhy stanovia podľa vzťahu:

$$v_j = f_j/m, \quad (1)$$

kde v_j je váha j -teho kritéria,

m je počet kritérií.

Prípad, keď počet preferencií určitého kritéria je nulový.

Ak aspoň v jednom prípade je celkový počet preferencií príslušného ukazovateľa rovný nule, musíme pre stanovenie štandardizovaných váh vypočítať najprv neštandardizovanú váhu j -teho ukazovateľa podľa vzťahu.

$$k_j = m + 1 - p_j \quad (2)$$

kde p_j - poradie j -teho ukazovateľa v jeho preferenčnom usporiadaní v komparačnej matici. Následne môžeme štandardizáciu váh jednotlivých ukazovateľov realizovať pomocou vzťahu:

$$v_j^* = \frac{k_j}{\sum_{j=1}^m k_j} \quad (3)$$

Normovanie váh potom uskutočníme podľa vzťahu:

$$v_j^{(m)} = k_j / \sum k_j. \quad (4)$$

Tabuľka 3 Stanovenie váh kritérií metódou párového porovnávania

Ukazovateľ	ROE	ROS _{op}	T/A	DSKZ	f_j	k_j	v_j
ROE	X	1	1	1	3	4	0,40
ROS _{op}	0	X	1	1	2	3	0,30
T/A	0	0	X	1	1	2	0,20
DSKZ	0	0	0	X	0	1	0,10
Spolu						10	1,00

Prameň: Vlastné spracovanie

Precíznejšiu verziu metódy párového porovnávania predstavuje tzv. *Saatyho metóda*.

Saatyho metóda pozostáva z dvoch krokov:

1. krok:

Zisťovanie preferenčných vzťahov dvojíc kritérií usporiadaných v tabuľke – na rozdiel matice párových porovnaní sa ale určuje aj veľkosť preferencie, ktorá sa vyjadruje určitým počtom bodov zo zvolenej bodovej stupnice

Počet bodov	Deskriptor
1	Kritériá sú rovnako významné
3	Prvé kritérium je slabo významnejšie
5	Prvé kritérium je dost' významnejšie
7	Prvé kritérium je preukázateľne významnejšie
9	Prvé kritérium je absolútne významnejšie

Hodnoty 2,4,6,8 sa dajú využiť k „jemnejšiemu“ rozlíšeniu veľkostí preferencií dvojíc kritérií

Výsledkom prvého kroku je získanie pravej hornej trojuholníkovej časti tzv. *matice veľkosti preferencií*, resp. *Saatyho matice*, resp. *matice relatívnych dôležitostí*.

Ak označíme túto maticu S , potom jej ďalšie prvky dostaneme podľa vzťahov:

$s_{ii} = 1$ pre všetky i ,

$s_{ji} = 1/s_{ij}$ pre všetky i a j .

2. krok:

Stanovenie váh jednotlivých kritérií pomocou Saatyho matice buď exaktnými alebo aproximatívnymi postupmi.

Exaktný – Saatyho postup: pomocou metódy najmenších štvorcov minimalizovať výraz

$$\hat{v}_j = \arg \min \sum_{i=1}^m \sum_{j=1}^m \left(s_{ij} - \frac{v_i}{v_j} \right)^2 \quad (5)$$

kde

v_i = (hľadaná) váha i -teho kritéria,

v_j = (hľadaná) váha j -teho kritéria,

pričom $\sum_{j=1}^m v_j = 1$, $v_i > 0$, $i = 1, 2, \dots, m$.

Pre úlohy menších rozmerov (napr. $m \leq 20$), môžeme úlohu riešiť pomocou funkcie *Riešiteľa* v programe *MS Excel*.

Pre nami vybrané ukazovatele sme zostavili nasledovnú Saatyho maticu:

Tabuľka 3 Saatyho matica

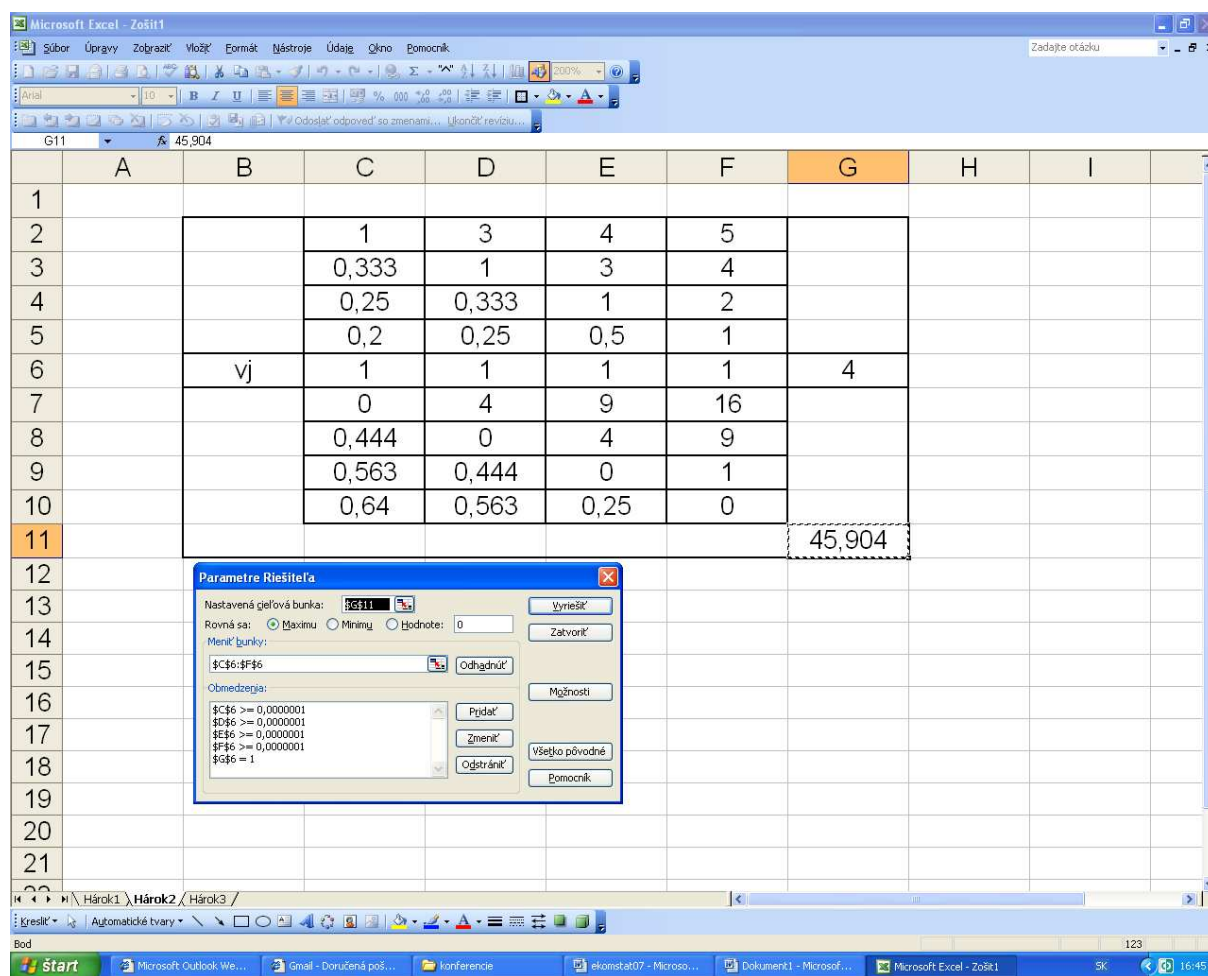
Ukazovateľ	ROE	ROS	T/A	DSKZ
ROE	1	3	4	5
ROS	1/3	1	3	4
T/A	1/4	1/3	1	2
DSKZ	1/5	1/4	1/2	1

Prameň: Vlastné spracovanie

Tabuľka 4 Východisková matica Saatyho metódy

	1,000	3,000	4,000	5,000	
	0,333	1,000	3,000	4,000	
	0,250	0,333	1,000	2,000	
	0,200	0,250	0,500	1,000	
vj	1,000	1,000	1,000	1,000	4,000
	0,000	4,000	9,000	16,000	
	0,444	0,000	4,000	9,000	
	0,563	0,444	0,000	1,000	
	0,640	0,563	0,250	0,000	
					45,904

Prameň: Vlastné spracovanie



Obrázok 1 Riešenie Saatyho metódy pomocou Riešiteľa v MS Excel

Tabuľka 5 Získané výsledky

	1,000	3,000	4,000	5,000	
	0,333	1,000	3,000	4,000	
	0,250	0,333	1,000	2,000	
	0,200	0,250	0,500	1,000	
v _j	0,478	0,315	0,120	0,087	1,000
	0,000	2,194	0,000	0,210	
	0,106	0,000	0,139	0,165	
	0,000	0,002	0,000	0,400	
	0,000	0,001	0,053	0,000	
					3,270

Prameň: Vlastné spracovanie

Aproximatívne môžeme pri Saatyho metódy použiť tzv. metódu logaritmických najmenších štvorcov (metóda geometrického priemeru).

Jej princíp spočíva v stanovení geometrických priemerov hodnôt riadkov Saatyho matice S . „Normalizáciou“ týchto riadkových priemerov (t.j. ich vydelením súčtom týchto geometrických priemerov) dostaneme odhady váh zodpovedajúcich kritérií.

Tabuľka 6 Určenie váh kritérií aproximatívnu Saatyho metódou

Ukazovateľ	ROE	ROS _{op}	T/A	DSKZ	Geometrický priemer	v _j
ROE	1	3	4	5	2,783	0,53
ROS _{op}	1/3	1	3	4	1,414	0,27
T/A	1/4	1/3	1	2	0,639	0,12
DSKZ	1/5	1/4	1/2	1	0,398	0,08
Spolu					5,234	1,00

Prameň: Vlastné spracovanie

3. Záver

Príspevok na reálnych dátach uviedol možnosť aplikácie vybraných metód stanovenia váh kritérií pri analyzovaní postavenia podniku na trhu. Zo získaných výsledkov zhrnutých v tabuľke 7 vidíme, že použitím variantných metód sme získali približne identické hodnoty stanovenia váh zvolených kritérií pre hodnotenie postavenia podniku na trhu.

Tabuľka 7 Zhrnutie výsledkov podľa použitej metódy

	SPAN	Komparačná matica	Exaktná Saatyho metóda	Aproximatívna Saatyho metóda
ROE	0,41	0,40	0,478	0,53
ROS	0,25	0,30	0,315	0,27
T/A	0,22	0,20	0,120	0,12
DSKZ	0,12	0,10	0,087	0,08

Prameň: Vlastné spracovanie

Problematika východísk realizácie analýzy postavenia podniku na trhu je samozrejme oveľa širšia (ako je uvedené aj v úvode tohto príspevku). Stanovenie váh kritérií je však jej neoddeliteľnou súčasťou.

4. Literatúra

FICZOVÁ, I. – SEDLÁČEK, J. – ÚRADNÍČEK, V. 2000. Finančná analýza podniku I. Banská Bystrica : RCPF, 2000. 174 s. ISBN 80-89029-02-7

FOTR, J. – DĚDINA, J. – HRŮZOVÁ H. 2003. Manažerské rozhodování. 3. vyd. Praha : Ekopress, 2003. 250 s. ISBN 80-86119-69-6

GAVLIAK, R. 2003. Tvorba hodnoty a metódy hodnotenia výkonnosti podnikov slovenskými bankami. In: Výkonnosť malých a stredných podnikov v Slovenskej republike a v Českej republike. Zborník príspevkov z medzinárodnej vedeckej konferencie. Banská Bystrica : Ekonomická fakulta UMB, 2003, s. 50 – 55. ISBN 80-8055-840-X.

KANDEROVÁ, M. 2003. Zhodnotenie vplyvu determinantov kapitálovej štruktúry na vzorke podnikov Slovenskej republiky. In : Ekonomie a management. Liberec : Hospodárska fakulta Technickej univerzity, č. 3/2003.

LUHA, J. 1993. *Metódy štatistickej analýzy kvalitatívnych znakov*. In: Zborník príspevkov Využitie štatistických metód v sociálno - ekonomickej praxi, EKOMSTAT'93. Trenčianske Teplice : SŠDS, 1993.

STANKOVIČOVÁ, I. 2006. *Základné princípy stratégie modelovania*. In: Forum Statisticum Slovaca 1/2006, s. 78 – 82. ISSN 1336-7420.

ÚRADNÍČEK, V. 1998. Alternatívne prístupy k hodnoteniu postavenia firmy na trhu v podmienkach Slovenskej republiky. In: Acta Facultatis Aeriae Publici 1/1998. Banská Bystrica : FF UMB, 1998, s. 89 – 102.

Adresa autora:

Vladimír Úradníček, Ing., PhD.
Ekonomická fakulta UMB
Katedra kvantitatívnych metód a informatiky
Oddelenie štatistiky a ekonomickej analytiky
Tajovského 10
975 90 Banská Bystrica
vladimir.uradnicek@umb.sk

Hlavné trendy vo vývoji ekonomiky SR

Vladimír Kvetan, Martin Mlýnek, Marek Radvanský, Viliam Páleník*

Abstract

The economy of the Slovak Republic is running through enormous success. More than 8% of GDP growth, which did not generated significant imbalances and unemployment rate close to 10% have to be evaluated positively. The expectations of future development are positive as well. The article is focused on presentation of midterm forecast of development of economy of SR. The forecast is based on econometric model ECM_BIER07Q1. This article is not only describing the forecast, but also key equations of the model. The key equations are equations for number of employed, monthly wage, household consumption, gross capital formation and equations of foreign trade. The presented forecast can be assigned as a optimistic scenario.

Key words: econometric modeling, forecast, error-correction model, Slovak Republic

Úvod

Vývoj ekonomiky SR zaznamenal v roku 2006 výrazné úspechy. Viac ako 8% rast HDP, ktorý negeneroval výrazné nerovnováhy, miera nezamestnanosti z disponibilného počtu mierne cez 10% nie je možné hodnotiť inak ako pozitívne. Nemenej pozitívne sú aj očakávania o budúcom vývoji. Predložený článok je zameraný prezentáciu strednodobej prognózy vývoja ekonomiky SR. Predkladaná prognóza vývoja ekonomiky SR vznikla na základe použitia ekonometrického modelu ECM-BIER07Q1¹, ktorý je autormi pravidelne modifikovaný a zdokonaľovaný minimálne raz ročne. Tento článok však sa neobmedzí iba na prezentáciu strednodobej prognózy, ale sú v ňom aj popísané kľúčové modelové rovnice, na základe ktorých táto prognóza vznikla.

Predpoklady prognózy

Prognóza je založená na predpokladoch o vývoji vonkajšieho a vnútorného ekonomického prostredia, ktoré vychádzajú z výsledkov ekonomiky dosiahnutých v roku

* Autori sú pracovníkmi Ekonomického ústavu SAV

¹ Vychádza z modelu ECM-ISWE06Q1

2006. Predpoklady prognózy takisto zohľadňujú očakávaný vývoj smerovania hospodárskej politiky vlády, centrálnej banky ako aj hospodársky vývoj Európskej únie a celej svetovej ekonomiky.

V otázke menovej politiky a plánovanej menovej integrácie predpokladáme, že vláda nepovolí v prípravách na vstup do eurozóny. Zavedenie eura preto predpokladáme na začiatok roka 2009, pričom konverzný kurz by sa mal pohybovať na spodnej hranici povoleného intervalu centrálnej parity. Snaha o spomenutý vstup do eurozóny predurčuje na základe inflačného kritéria predpoklad o cenovom vývoji. Pri vývoji regulovaných cien preto uvažujeme opatrnú a rozumnú politiku Úradu pre reguláciu sieťových odvetví, ktorá nebude vnášať do cenového vývoja výrazné inflačné tlaky. Z toho je možné prijať predpoklad, že pomer jadrovej a celkovej inflácie bude celkovo vyrovnaný. S podmienkami Maastrichtských kritérií bude zákonite korešpondovať aj vývoj deficitu verejných financií. Predpokladáme, že podiel schodku štátneho rozpočtu by mal dosiahnuť úroveň pod 3% HDP už v tomto roku (2007).

Pre vývoj na trhu práce je predurčujúci vývoj dopytu a ponuky práce. Ponuka práce je závislá najmä od demografického vývoju. Pri predikcii demografického vývoja bol použitý stredný variant Výskumného demografického centra v Bratislave s predpokladom mierneho rastu miery participácie. Takisto predpokladáme, že prezentovaný model očakávanej zmeny zákonníka práce mierne nepriaznivo ovplyvní vývoj trhu práce po roku 2007, predovšetkým znížením dynamiky tvorby nových pracovných miest. Vzhľadom na ostatné pozitívne trendy vývoja však nepredpokladáme výrazný vplyv na dynamiku rastu ekonomiky v strednodobom horizonte.

Pri vývoji vonkajšieho prostredia je zásadným predpokladom pretrvávajúci vyvážený rast reálneho HDP v EU 15 na úrovni 2%. Z toho vychádza náš predpoklad o raste vonkajšieho prostredia, kde je predpokladaný rast importu EU15 vo výške približne 6%. Takýto vývoj vonkajšieho prostredia je určujúci aj pre vývoj prílevu zahraničného kapitálu. Priame zahraničné investície sú v horizonte prognózy predpokladané v priemere 1,65 mld. USD ročne. Závažným problémom pri zostavovaní prognózy bola otázka Národného strategického referenčného rámca, ktorého schválenie a implementácia je stále nejasná.

Prognóza HDP vychádza z vývoja domáceho dopytu a čistého exportu, pričom hodnotíme ich najvýznamnejšie determinanty (zamestnanosť, mzda, kurz, zahraničný dopyt, PZI a pod.).

Kľúčové modelové vzťahy

Z modelového hľadiska sú kľúčovými rovnicami vnútorného dopytu rovnice pre výpočet priemerného počtu zamestnaných, výšku nominálnej mzdy, konečnej spotreby domácností a tvorby hrubého fixného kapitálu. Tieto rovnice (rovnako ako všetky rovnice v modeli) boli odhadnuté na základe štvrťročných údajov z časového obdobia 1995q1 až 2006q4 pri použití metódy korekčného člena (ECM)

Rovnica (1) prezentuje závislosť priemerného počtu zamestnaných podľa VZPO *L1* od ceny práce *W/PPI00* a HDP *Y*.

$$DLOG(L1) = 1.908 - 0.153*(LOG(L1(-1)) + 1.03*LOG(W(-1)/PPI00(-1)) - 0.888*LOG(Y(-1))) - 0.0571*DLOG(W/PPI00) + 0.0541*U95Q1 + 0.0165*UL1 \quad (1)$$

Je zrejmé, že z dlhodobého hľadiska je dopyt po práci ovplyvňovaný cenou práce negatívne, nakoľko platí zákon dopytu a ceny a pozitívne vývojom HDP kde pri raste produkcie je zamestnávateľ schopný a ochotný zamestnávať viac ľudí. Z krátkodobého hľadiska je dopyt po práci ovplyvňovaný iba cenou práce, kde pri raste ceny práce klesá zákonite dopyt po nej.

Priemerná mesačná mzda *W* (rovnica (2)) je v modeli determinovaná mierou nezamestnanosti *UR1* a cenovým vývojom *CPI00*.

$$DLOG(W) = 1.706 - 0.167*(LOG(W(-1)) + 0.281*LOG(UR1(-1)) - 1.089*LOG(CPI00(-1))) + 0.130*(@SEAS(4)) - 0.131*(@SEAS(1)) + 0.0639*(@SEAS(2)) + 0.0359UW \quad (2)$$

Mzda je spomenutými faktormi ovplyvňovaná najmä z dlhodobého hľadiska, nakoľko táto je indexovaná predovšetkým vzhľadom na historický vývoj. Je zrejmé že rast miery nezamestnanosti pôsobí na mzdový vývoj negatívne. Platí zákon, že čím je neuspokojená ponuka práce vyššia, tým je na trhu viac ľudí ochotných pracovať za nižšiu mzdu a teda je vyvíjaný menší tlak na jej rast. Na druhej strane rast životných nákladov meraný indexom spotrebiteľských cien tlačí zamestnancov aby žiadali od zamestnávateľov vyššie mzdy. Z krátkodobého hľadiska je mzda ovplyvňovaná najmä exogénnymi zásahmi zo strany zamestnávateľov a vlády. Rast mzdy je napríklad charakteristický v štvrtom štvrťroku, kde sú vyplácané koncoročné odmeny, trináste a štrnáste platy a pod.

Mzda a priemerná zamestnanosť sa v modeli následne agregujú do pracovných príjmov obyvateľstva z ktorých sú ďalej odvodené zložky hrubého disponibilného dôchodku domácností tak, ako sú evidované v národnom účtovníctve podľa metodiky ESA 95. Hrubý

disponibilný dôchodok *HDD* je spolu s výškou úverov poskytnutým domácnostiam *ACH* modelovým determinantom konečnej spotreby obyvateľstva *HCP* (rovnica (3))

$$DLOG(HCP) = 0.0019 - 0.398*(LOG(HCP(-1)) - 0.961*LOG(HDD(-1)) - 0.0331*LOG(ACH3B(-1))) + 0.349*DLOG(HDD) + 0.0303*DLOG(ACH3B) - 0.034*UC \quad (3)$$

Je samozrejmé, že obyvateľstvo svoj disponibilný príjem na svoju spotrebu, pričom nespotrebovaná časť disponibilného príjmu je zdrojom úspor. V ostatnom čase sa však na spotrebe domácností významnou mierou podieľajú aj úvery poskytované domácnostiam bankami. Najmä v ostatnom období sa banky orientujú na poskytovanie spotrebných úverov obyvateľstvu. Je to možné odôvodniť nižším rizikovým stupňom takýchto klientov ako aj celkovo nižšou priemernou výškou úveru.

Ďalšou významnou zložkou vnútorného dopytu je tvorba hrubého fixného kapitálu *DFKP*, ktorá prezentuje objem investícií. V modeli ECM-BIER07Q1 sú jej hlavnými determinantmi výška poskytnutá podnikom a domácnostiam *ACEH* a výška vlastných príjmov, ktorá je odvodená od zisku *PROF* zníženého o platené dane z príjmu právnických osôb *TYZ*. V neposlednom rade je pre rast investícií určujúci vývoj kapitálových výdavkov štátneho rozpočtu *SBCE*.

$$DLOG(DFKP) = 0.0301 - 0.689*(LOG(DFKP(-1)) - 0.311*LOG(ACEH(-1)) - 0.564*LOG(PROF(-1) - TYZ(-1)) - 0.111*LOG(SBCE(-1))) + 0.0733*DLOG(SBCE) - 0.51*DLOG(ACEH) - 0.164*@SEAS(3) + 0.0445*@SEAS(2) - 0.091*UDFKP \quad (4)$$

Je zrejmé, že všetky vysvetľujúce premenné predstavujú zdroje investícií a pôsobia na rast fixného kapitálu kladne. Neprehliadnuteľným zdrojom financovania investícií budú aj naďalej priame zahraničné investície a príjmy z kohézneho a štrukturálnych fondov EÚ. Nakoľko podstata ekonometrických rovníc je v odhade budúcnosti na základe minulosti, z dôvodu výrazne odlišného vývoja PZI a fondov EÚ sú tieto vplyvy zahrnuté v exogénnej premennej *UDFKP*.

Popri domácom dopyte je z dôvodu vysokej otvorenosti ekonomiky kľúčový vývoj zahraničného obchodu. Vstupom SR do EÚ ako aj prílevom PZI, ktorý je zaznamenávaný v ostatnom období bude otvorenosť ekonomiky ako aj význam zahraničného obchodu narastať. Preto aj v modeli ECM-BIER07Q1 sú rovnice pre výpočet zahraničného obchodu rovnako dôležité ako rovnice pre výpočet zložiek vnútorného dopytu.

Hlavnými faktormi, ktoré v modeli ovplyvňujú výšku exportu *EGSR* je HDP EÚ15 *Y_EU15*, kurz SKK k EUR *EURSKP* a pomer medzi cenami dovozu EÚ15 *PIM_EU15* a deflátorom slovenského vývozu *PEGSR* (rovnica (5)).

$$DLOG(EGSR) = -7.382 - 0.205*(LOG(EGSR(-1)) - 5.363*LOG(Y_EU15(-1))) + 1.442*DLOG(Y_EU15) + 0.74*DLOG(EURSKP) + 0.429*DLOG(PIM_EU15/PEGSR) + 0.0786*(@SEAS(2)) - 0.0386*UEXP \quad (5)$$

HDP EU 15 charakterizuje svetový dopyt po slovenskom exporte, nakoľko tieto krajiny tvoria najvýznamnejších vývozných partnerov SR. Zhodnocovanie kurzu SKK k EUR pochopiteľne pôsobí proti rastu exportu. Tento vzťah je však možné sledovať z dlhodobého hľadiska. V súčasnom období sa deje, že export narastá aj pri zhodnocovaní koruny. Túto anomáliu vymykajúcu sa zo štandardných teoretických záverov je však možné vysvetliť zmenou štruktúry nášho exportu.

Spomenutá zmena štruktúry exportu bude vplývať na rast jeho dovoznej náročnosti. Rovnako nepredpokladáme výrazný pokles dovoznej náročnosti domáceho dopytu *DD*. Objem exportu bude teda spolu s výškou vnútorného dopytu hlavnými určujúcimi faktormi objemu dovozu *MGRS*.

$$DLOG(MGRS) = -0.135 - 0.3*(LOG(MGRS(-1)) - 0.244*LOG(DD(-1)) - 0.829*LOG(EGSR(-1))) + 0.587*DLOG(DD) + 0.734*DLOG(EGSR) + 0.0756*UIM \quad (6)$$

Prognóza vývoja ekonomiky SR

Z uvedených kľúčových rovníc, ktoré sú začlenené do modelu a predpokladov pretransformovaných do exogénnych premenných bola vypočítaná strednodobá prognóza ekonomiky SR. Celkovo sa táto prognóza dá hodnotiť ako optimistická.

V rokoch 2007 a 2008 bude rast HDP založený na raste domáceho dopytu. Na rozdiel od minulosti nebude rast domáceho dopytu výrazne znižovaný negatívnym saldom zahraničného obchodu. Naopak, takmer vyrovnaná bilancia zahraničného obchodu s tovarmi a službami bude pôsobiť na HDP neutrálne.

Najvýznamnejšími faktormi určujúcimi vývoj domáceho dopytu sú mzda a zamestnanosť, objem úverov poskytnutých domácnostiam (pre konečnú spotrebu domácností); objem úverov, kapitálové výdavky štátneho rozpočtu, prílev zahraničných investícií (pre vývoj tvorby hrubého fixného kapitálu).

Priemerný 1,2 %-ný rast evidovanej zamestnanosti pre obdobie 2007-2015 vyplýva zo stabilného rastu ekonomiky. Z titulu vyrovnaného demografického vývoja² miera nezamestnanosti počítanej na základe VZPS do roku 2009 poklesne na 9,5 %. Pozitívny ekonomický vývoj sa prenesie aj do vývoja nominálnej mzdy. Jej rast sa bude však postupne

² Nepredpokladáme žiadne výrazné zásahy, ktoré by mohli viesť poruchu do demografických trendov a ani do výraznejšej zmeny miery participácie

spomašovať z 8,8% v roku 2007 na 5,8 v roku 2015. Pri miernej inflácii (znižujúci trend z 3,4 v roku 2007 do 2,2 v roku 2015) bude reálna mzda v roku 2007 rásť rýchlosťou 5,4%. Rastúca mzda a zamestnanosť pri súčasnom raste úverov domácnostiam a poklese úrokových mier z vkladov umožnia rast reálnej konečnej spotreby domácností v roku 2007 o 7,5%. V nasledujúcich rokoch sa rast zamestnanosti, priaznivý vývoj cien a miezd pozitívne odrazí na vývoji konečnej spotreby domácností, ktorej rast by mal dosahovať v rokoch 2007 – 2015 priemerne 6%. Pre vývoj tvorby hrubého fixného kapitálu bude určujúci očakávaný rast outputu a nastúpený trend prílevu zahraničných investícií. Súčasný vývoj zatiaľ nepotvrďuje obavy, ktoré vnášali do investičného vývoja predvolebné prehlásenia a očakávané nové vládne zoskupenie. Z tohto dôvodu na rok 2007 predpokladáme opätovné zvýšenie dynamiky rastu hrubého fixného kapitálu na úroveň 13,7%. Počas celého horizontu prognózy predpovedáme rast investícií v priemere 9% ročne.

Podľa ekonomickej teórie majú byť hlavnými faktormi pôsobiacimi na vývoz vývoj reálneho výmenného kurzu a vývoj zahraničného dopytu. Keďže slovenská ekonomika a štruktúra jej exportu prechádza výraznými zmenami, čistý vývoz sa v uplynulých rokoch vyvíjal mimoriadne priaznivo aj napriek pokračujúcemu zhodnocovaniu reálneho výmenného kurzu a neutrálnemu vývoju vonkajšieho prostredia. Podobne ako v uplynulom období aj v najbližších rokoch budú mať na vývoj exportov najvýraznejší vplyv priame zahraničné investície³. Pre vývoj vonkajšieho dopytu bol pri prognóze zohľadnený očakávaný mierny hospodársky rast krajín EÚ15. Z týchto dôvodov očakávame aj pri pretrvávajúcom reálnom zhodnocovaní kurzu rast vývozu tovarov a služieb v stálych cenách v roku 2007 o 15,9%. V ďalšom vývoji bude dynamika dovozu postupne spomaľovať na 10,7% v roku 2015. Dovoz je determinovaný vnútorným dopytom a exportom. Smerom k rastu dovozu bude pôsobiť obnovenie rastu domáceho dopytu, spotrebného a investičného charakteru. V roku 2007 vzrastie dovoz podľa našich odhadov o 15,1%.

Pôsobenie spomenutých faktorov bolo spojené do výslednej prognózy HDP. Vplyvom oživenia rastu domáceho dopytu a nižšieho rastu dovozu sa do určitej miery zníži čistý vývoz, čo v roku 2007 povedie k medziročnému reálnemu rastu HDP o 9,4 %. Do roku 2015 prognózuujeme priemerné ročné tempo rastu ekonomiky na úrovni 7,5 %.

³ Podrobnejšiu analýzu vplyvu priamych zahraničných investícií na exportnú výkonnosť SR podáva Tóth a Chudík (2004)

Záver

Predložená prognóza vznikla na základe modelového prístupu. Bol použitý model ECM_BIER07Q1. Tento model pozostáva s radu rovníc a identít, z ktorých sú kľúčové rovnice pre počet zamestnaných, nominálnu mzdu, zložky vnútorného dopytu a zahraničného obchodu.

Predložená prognóza predstavuje optimistický variant vývoja ekonomiky SR. Nepredpokladá takú zmenu smerovania politiky, ktorá by výraznejšie tlmila naštartovanú prosperitu a rast. Rovnako predpokladáme, že určujúcim bude snaha vlády SR o vstup do EMU v roku 2009.

Prognóza je charakteristické výrazným rastom HDP v najbližšom období, pričom rast by sa zo strednodobého hľadiska mal stabilizovať na úrovni cez 6%. To bude mať za následok výrazné zlepšenie situácie na trhu práce. Na základe predloženej modelovej prognózy je preto možné očakávať pokles miery nezamestnanosti až na úroveň prirodzenej infláciu negenerujúcej miery nezamestnanosti.

Tabuľka 1 Prognóza hlavných makroekonomických ukazovateľov

	Skutočnosť				Prognóza, tempá rastu									
	2004	2005	2006	tempo rastu	2007	2008	2009	2010	2011	2012	2013	2014	2015	Priemer 2007-15
HDP, s.c.	1110.8	1177.9	1275.3	8.2	9.4	8.5	8.2	8.0	7.4	6.9	6.5	6.1	6.1	7.5
Konečná spotreba domácností, s.c.	603.9	647.3	688.1	6.3	7.5	7.3	6.9	6.3	5.5	5.1	5.0	5.0	4.9	6.0
Konečná spotreba verejnej správy, s.c.	220.2	218.9	227.8	4.1	7.3	5.4	6	6.1	5.8	5.3	5.4	5.2	5.3	5.8
Tvorba hrubého fixného kapitálu, s.c.	281.5	330.7	355	7.4	13,7	11.9	11.8	10.1	10.5	8.3	6.9	5.3	6.3	9.4
Vývoz výrobkov a služieb, s.c.	925.2	1052.5	1270.5	20.7	16.0	12.6	13.7	14.4	12.5	11.9	11.3	10.9	10.7	
Dovoz výrobkov a služieb, s.c.	952.3	1110.5	1308.2	17.8	15.1	12.0	13.7	13.8	12.1	11.2	10.7	10.2	10.2	
Index cien priemyselných výrobcov	1.23	1.29	1.39	7.8	4.1	3.2	3.1	3	2.6	2.5	2.4	2.4	2.4	2.9
Index spotrebiteľských cien	1.29	1.33	1.4	5.3	3.4	3.3	2.8	2.4	2.5	2.5	2.4	2.2	2.2	2.6
Priemerná mesačná mzda, nominálna, Sk	15817	17260	18761	8.7	8.8	8.5	8.2	6.4	6.2	5.9	5.9	5.8	5.8	6.8
Priemerná mesačná mzda, reálna, Sk	12212	12971	13401	3.3	5.4	5,0	5.2	3.9	3.7	3.4	3.4	3.4	3.5	4.1
Zamestnanosť, tis. osôb, VZPS	2170.4	2216.2	2301.4	3.8	2.1	2,0	1.6	1.4	1.2	0.8	0.6	0.4	0.2	
absolútne hodnoty														
Saldo zahraničného obchodu, s.c.	-27.1	-58.1	-37.7		-32.7	-28	-31.7	-23.3	-18	-3.4	14.1	36.6	55.6	
Nezamestnanosť, v %, VZPS	18.1	16.2	13.3	-	11.7	10.4	9.5	8.7	8	7.4	6.9	6.6	6.3	
Príjmy z DPH, b.c	99.6	122.4	128.7		138.2	150.8	162.2	173.5	185.2	196.2	207	217.6	229.2	
Príjmy zo spoterbných daní, b.c.	43.4	50	52.1		57.1	62.7	68.5	74.2	79.5	84.9	90.3	96	101.9	

Literatúra:

- [1] Aktualizovaný konvergenčný program Slovenska na roky 2004 až 2010, Ministerstvo financií SR, Bratislava, 2004
- [2] Menový program NBS do roku 2008, Národná banka Slovenska, Bratislava, 2004
- [3] Prognóza vývoja obyvateľstva SR do roku 2050, Výskumné demografické centrum, Bratislava, 2002
- [4] European Commission Spring 2005 Economic Forecasts, European Commission, 2005
- [5] Chajdiak, J.: Ekonomický vývoj SR. In. Pohľady na ekonomiku Slovenska 2004. Bratislava 6.4.2004. Slovenská štatistická a demografická spoločnosť, 2004. ISBN 80-88946-33-6. s. 9–16.
- [6] Haluška, J. - Olexa, M.: Vybrané výsledky makroekonomickej prognózy slovenskej ekonomiky v roku 2006 s výhľadom do roku 2009. In: Forum Statisticum Slovaca 2/2006 ISSN 1336 -7420.
- [7] Tóth, J. - Chudík, A.: Slovak Foreign Trade Model., ING Economics Paper No. 3., ING Bank, Bratislava, 2004

Miera zamestnanosti ako štruktúrny ukazovateľ lisabonskej stratégie

Viliam Páleník, Martin Mlýnek¹

Abstrakt: Following paper aims on Mr. Wim Kok report of fulfilling aims of Lisbon strategy. Aims of Lisbon strategy were critically reviewed after Mr. Wim Kok report and changed. Nevertheless some of the aims of strategy seem to be infeasible in respect of development of EU economies. This paper discuss feasibility of labour market aims of strategy in respect to development of Slovak labour market.

Key words: Lisbon strategy, labour market, structural indicators

V marci roku 2000 sa lídri krajín Európskej Únie zhodli na stratégii rozvoja EU do roku 2010 s cieľom vytvoriť „Najdynamickejšiu a konkurencie schopnú vedomostnú ekonomiku na svete, s udržateľným rastom a zamestnanosťou, vyššou sociálnou kohéziou s rešpektom k životnému prostrediu“. Lisabonská stratégia, ako vošla táto stratégia na známosť, je formulácia komplexnej série vzájomne závislých reformných krokov. Splnenie vytýčeného cieľa podľa nej vyžaduje urýchlené konanie v piatich oblastiach politiky a to vytvorenie (1) „vedomostnej spoločnosti“ zvyšovaním atraktívnosti Európy pre vedcov, predovšetkým zameraním sa na vedu a výskum ako hlavnú prioritu, (2) dokončenie tvorby spoločného trhu pre voľný pohyb tovarov a kapitálu a neodkladne pracovať na jednotnom trhu služieb. (3) Znížiť celkové administratívne zaťaženie podnikateľov, zlepšiť kvalitu právnych predpisov, uľahčiť rozbehnutie nových podnikov a tým výrazne zlepšiť podnikateľskú klímu. (4) Na trhu práce vypracovať stratégie celoživotného vzdelávania a aktívneho starnutia a konať podľa odporúčaní európskej skupiny odborníkov na zamestnanosť, a taktiež v oblasti (5) životného prostredia šíriť ekologické novinky a vytvárať si náskok pri rozvoji eko-priemyslu. Pri uskutočňovaní týchto politík by mali členské krajiny kooperovať politiky a riadiť sa poznatkom, že konanie ktoréhokoľvek členského štátu bude omnoho efektívnejšie, ak budú ostatné členské štáty konať rovnako.

1. Správa Wima Koka

V marci 2004 Bruselské zasadanie Európskej rady vyzvalo Európsku komisu, aby vytvorila pracovnú skupinu na vysokej úrovni pod vedením pána Wima Koka na vytvorenie nezávislej revízie Lisabonskej stratégie, ktorá prispeje k priebežnej hodnotiacej správe. Správa pracovnej skupiny mala identifikovať opatrenia, ktoré spolu tvoria dôslednú stratégiu pre európske ekonomiky na dosiahnutie zámerov a cieľov formulovaných v Lisabone. Pracovnú skupinu mal tvoriť obmedzený počet odborníkov schopných vyjadriť názory všetkých zainteresovaných strán. Správa sa mala uverejniť a predložiť Komisii do 1. novembra 2004. Pra-

¹ Adresa autorov: doc. RNDr. Viliam Páleník, PhD., Ing. Martin Mlýnek, Ekonomický ústav SAV, Šancová 56, 811 05
tel: +421 2 5249 7053, kl. 140 fax: +421 2 5249 5106
e-mail: viliam.palenik@savba.sk, matin.mlynek@savba.sk, www.ekonom.sav.sk

covná skupina na vysokej úrovni vykonávala svoju činnosť od mája do októbra 2004. Zišla sa šesť krát a 3. novembra 2004 predložila Komisii svoju správu.²

V Lisabone v roku 2000 a na následných jarných zasadaniach Európskej rady vznikla séria ambiciózných cieľov na podporu rozvoja európskeho hospodárstva. Na monitorovanie pokroku lisabonskej stratégie sa Komisia a Rada dohodli na zozname 14 ukazovateľov. Základné ciele formulované v Lisabone sú dosiahnutie miery zamestnanosti v Európskej únii na úrovni 67 % do roku 2005 a 70 % do roku 2010, taktiež dosiahnutie miery zamestnanosti žien na úrovni 57 % do roku 2005 a 60 % do roku 2010 ako aj miery zamestnanosti starších pracovníkov na úrovni 50 % do roku 2010. Ďalšie ciele sú dosiahnutie výdavkov 3 % z HDP na výskum a vývoj do roku 2010 a zníženie emisii skleníkových plynov na úroveň 92 % oproti roku 1990 do roku 2010 v rámci dodržiavania Kjótskeho protokolu. Podľa správy Wim Koka polceste k roku 2010 bol celkový dojem veľmi zmiešaný a mnohým krajinám sa nedarilo uskutočňovať formulované ciele. Prijatím 10 nových členských krajín do európskej únie ostáva splnenie týchto cieľov ešte viac otázne. Nové členské krajiny v mnohých oblastiach zaostávajú a to hlavne v oblasti hospodárskej výkonnosti ich ekonomík meranej cez hrubý domáci produkt na obyvateľa v parite kúpnej sily, ako aj v produktivite práce. V oblasti zamestnanosti mnohé nové členské krajiny výrazne zaostávajú a to hlavne vďaka dlhodobej nezamestnanosti. Výdavky na výskum a vývoj u nich tvoria zanedbateľný podiel z HDP. Jediný z cieľov, ktorý mnohé nové členské krajiny spĺňajú je zníženie emisie skleníkových plynov pod cieľovú úroveň.

2. Štruktúrne ukazovatele lisabonskej stratégie

S lisabonským procesom je spojených viac než sto ukazovateľov, podľa ktorých bude zrejme každá krajina v tom či onom ukazovateli hodnotená ako najlepšia. Jedno z kľúčových odporúčaní správy Wima Koka bolo: „Európska komisia by mala na jarné zasadanie Európskej rady dodať čo najvernejšiu výročnú „ligovú tabuľku“ pokroku členských štátov v dosahovaní štrnástich kľúčových ukazovateľov a cieľov. Krajiny, ktoré si dobre počínali, by mali byť ocenené; tie, ktoré sa neusilovali, treba zahanbiť.“ V takejto jednoznačnej podobe sa toto doporučenie nebolo prijaté a zavedené do praxe, avšak napriek tomu alebo práve preto stojí za pozornosť. Konkrétne hodnoty sú uvedené v Tabuľke 1.

² Pracovná skupina mala toto zloženie:

1. Wim Kok (predseda), bývalý predseda vlády Holandska;
2. Romain Bausch, prezident a výkonný riaditeľ firmy SES Global (Luxemburg);
3. Niall FitzGerald, predseda predstavenstva firmy Reuters, predsedajúci dialógu Trans- Atlantic Business;
4. Antonio Gutiérrez Vegara, člen španielskeho parlamentu;
5. Will Hutton (hovorca), konateľ nadácie Work Foundation;
6. Anne-Marie Idrac, predsedníčka Régie autonome des transports parisiens (RATP);
7. Wanja Lundby-Wedin, predsedníčka švédskej Konfederácie odborových zväzov (LÖ);
8. Thomas Mirow, bývalý člen mestskej rady Hamburgu, vyšší obchodný poradca;
9. Bedřich Moldan, predseda Strediska životného prostredia (Karlova univerzita, Praha);
10. Luigi Paganetto, profesor medzinárodnej ekonómie (Univerzita Tor Vergata, Rím);
11. Dariusz Rosati, profesor ekonómie, od júna 2004 člen Európskeho parlamentu ;
12. Veli Sundbäck, viceprezident firmy Nokia, Fínsko;
13. Friedrich Verzetnitsch, predseda Rakúskej federácie odborových zväzov (ÖGB), člen rakúskeho parlamentu.

Tabuľka 1: Relatívne údaje štruktúrnych ukazovateľov lisabonskej stratégie

		EU - 15															
Ukazovateľ		AT	BE	DE	DK	ES	FI	FR	EL	IE	IT	LU	NL	PT	SE	UK	
HDP na obyvateľa v parite kúpnej sily	2003	111	107	99	113	87	101	104	73	122	98	195	110	68	106	109	
Produktivita práce na zamestnaného v parite kúpnej sily	2003	96	118	94	98	95	99	114	90	120	104	132	95	64	97	102	
Miera zamestnanosti (v %)	2003	69	60	65	75	60	68	63	58	65	56	63	74	67	73	72	
Miera zamestnanosti žien (v %)	2003	63	52	59	71	46	66	57	44	56	43	52	66	61	72	65	
Miera zamestnanosti starších pracovníkov (v %)	2003	30	28	40	60	41	50	37	42	49	30	30	45	51	69	56	
Dosiahnutý (stredoškolský) stupeň vzdelanosti mladých ľudí vo veku 20-24 rokov	2003	84	81	73	74	63	85	81	82	86	70	70	73	48	86	78	
Výdaje na výskum a rozvoj ako percentuálny podiel na HDP	2003	2	2	3	3	1	3	2	1	1	1	2	2	1	4	2	
Investície (v súkromnom sektore) ako percentuálny podiel na HDP	2003	20	18	16	18	22	15	16	22	20	17	15	17	19	13	15	
Porovnanie cenovej úrovne s úrovňou v EU-15	2002	102	99	104	131	82	123	100	80	118	95	100	102	74	117	108	
Miera riziku chudoby	2003	12	13	11	10	19	11	15	20	21	19	12	11	20	9	17	
Dlhodobá miera nezamestnanosti	2003	1	4	5	1	4	2	4	5	2	5	1	1	2	1	1	
Regionálny rozptyl miery nezamestnanosti	2003	3	8	6	-	9	6	5	4	-	17	-	2	4	4	6	
Emisie skleníkových plynov	2002	109	102	81	99	139	107	98	127	129	109	85	101	141	96	85	
Celková vnútroštátna spotreba energie ako podiel na HDP	2002	146	214	165	123	229	272	187	258	164	184	198	202	254	224	212	
Množstvo nákladnej prepravy pomerne ku HDP	2002	120	100	102	85	137	95	96	127	133	103	110	97	126	90	86	

Tabuľka 1 (pokračovanie): Relatívne údaje štrukturálnych ukazovateľov lisabonskej stratégie

		EU -10															
																Cieľ	
Ukazovateľ		CY	CZ	EE	HU	LT	LV	MT	PL	SI	SK	EU-25	EU-15	US		2005	2010
HDP na obyvateľa v parite kúpnej sily	2003	76	67	43	56	42	39	68	42	71	47	91	100	140			
Produktivita práce na zamestnaného v parite kúpnej sily	2003	77	61	43	63	44	40	82	50	70	54	93	100	122			
Miera zamestnanosti (v %)	2003	69	65	63	57	61	62	54	51	63	58	63	64	71		67	70
Miera zamestnanosti žien (v %)	2003	60	56	59	51	58	58	34	46	58	52	55	56	66		57	60
Miera zamestnanosti starších pracovníkov (v %)	2003	50	42	52	29	45	44	33	27	24	25	40	42	60			50
Dosiahnutý (stredoškolský) stupeň vzdelanosti mladých ľudí vo veku 20-24 rokov	2003	82	92	81	85	82	74	43	89	91	94	77	74				
Výdaje na výskum a rozvoj ako percentuálny podiel na HDP	2003	0	1	1	1	1	0		1	2	1	2	2	3			3
Investície (v súkromnom sektore) ako percentuálny podiel na HDP	2003	14	22	25	20	18	23		15	21	23	17	17				
Porovnanie cenovej úrovne s úrovňou v EU-15	2002	83	53	61	55	51	54	72	58	73	44	96	100	113			
Miera risku chudoby	2003	16	8	18	10	17	16	15	15	11	21	15	15				
Dlhodobá miera nezamestnanosti	2003	1	4	5	2	6	4	4	11	3	11	4	3				
Regionálny rozptyl miery nezamestnanosti	2003	-	6	-	9	-	-	-	7	-	8	13	12				
Emisie skleníkových plynov	2002	150	74	45	69	40	37	129	68	99	72	91	97	113			92
Celková vnútroštátna spotreba energie ako podiel na HDP	2002	280	921	1156	564	1273	759	264	650	343	964	210	191	330			
Množstvo nákladnej prepravy pomerne ku HDP	2002	93	100	177	91	119	123		70	92	62	101	102	91			

V tabuľke podľa jednotlivých ukazovateľov vidieť postavenie jednotlivých krajín navzájom, ale aj vo vzťahu k priemeru EU25, EU 15, Spojeným štátom americkým a cieľom lisabonskej stratégie k roku 2005 a 2010. Definície použitých ukazovateľov sú v prílohe. Z hľadiska tejto práce je dôležité postavenie Slovenska, ktoré je prehľadne zosumarizované v Tabuľke 2. Ukazovatele sú zoradené od tých, kde má Slovensko najlepšiu pozíciu po tie, kde je jeho situácia najhoršia.

Tabuľka 2: Postavenie Slovenska podľa štruktúrnych ukazovateľov lisabonskej stratégie (zoradené od najlepšieho po najhoršie)

	Structural Indicator	Štruktúrny ukazovateľ
1.	Educational attainment (20-24) (%)	Dosiahnutý (stredoškolský) stupeň vzdelanosti mladých ľudí vo veku 20-24 rokov
2.	Business investment (% GDP)	Investície (v súkromnom sektore) ako percentuálny podiel na HDP
3.	Volume of transport	Množstvo nákladnej prepravy pomerne ku HDP
4.	Greenhouse gas emissions (index base year = 100)	Emisie skleníkových plynov
5.	Employment rate females (%)	Miera zamestnanosti žien (v %)
6.	GDP per capita (PPS, EU-15 = 100)	HDP na obyvateľa v parite kúpnej sily
7.	Labour productivity per person employed (PPS, EU-15 = 100)	Produktivita práce na zamestnaného v parite kúpnej sily
8.	Employment rate (%)	Miera zamestnanosti (v %)
9.	Employment rate of older workers (%)	Miera zamestnanosti starších pracovníkov (v %)
10.	Dispersion of regional employment rates	Regionálny rozptyl miery nezamestnanosti
11.	Energy intensity of the economy	Celková vnútroštátna spotreba energie ako podiel na HDP
12.	Research and development expenditure (% GDP)	Výdaje na výskum a rozvoj ako percentuálny podiel na HDP
13.	Comparative price levels (EU-15 = 100)	Porovnanie cenovej úrovne s úrovňou v EU-15
14.	Long-term unemployment rate (%)	Dlhodobá miera nezamestnanosti
15.	At-risk-of-poverty rate (%)	Miera riziku chudoby

Bližší pohľad na pozíciu Slovenska odhalí značne nevyváženosti slovenskej ekonomiky. Na jednej strane je Slovensko s niektorými ukazovateľmi na špici v rámci Európskej únie a na druhej je hlboko pod priemerom EU. Toto je dôsledok premeny slovenskej ekonomiky z centrálne riadenej na trhovú ekonomiku, ako aj priorit jednotlivých vlád a ich politík. Koncentrácia týchto faktorov ma za následok pretrvávajúcu značnú rozdielnosť slovenskej ekonomiky od ekonomík EU15. Slovensko navyše zaostáva v porovnateľnej cenovej hladine, v miere dlhodobej nezamestnanosti a v miere rizika chudoby.

3. Miera zamestnanosti ako štruktúrne ukazovatele lisabonskej stratégie

Mierou zamestnanosti je podiel zamestnaných na celkovej populácii v rovnakej vekovej skupine, v tomto (prípade 15 až 65 rokov). Nízka miera zamestnanosti na signalizuje nízke využitie ekonomického potenciálu krajiny s nepriamo aj požiadavky na verejné alebo súkromné zdroje, ktoré majú solidárne zabezpečenie nepracujúcich občanov. Cieľom lisabonskej stratégie je 70 % miera zamestnanosti odpovedajúca skutočnosti v USA (Tabuľka 3). Priemer EU25 je 63 %. Tento cieľ už spĺňa Dánsko, Holandsko a Švédsko. Slovenská miera zamestnanosti je 58 %, horšie je na tom len Malta a Poľsko.

Tabuľka 3: Miera zamestnanosti ako štruktúrny ukazovateľ (15 až 65 ročný)

		EU - 15														
Ukazovateľ		AT	BE	DE	DK	ES	FI	FR	EL	IE	IT	LU	NL	PT	SE	UK
Miera zamestnanosti (v %)	2003	69	60	65	75	60	68	63	58	65	56	63	74	67	73	72

		EU -10														Cieľ	
Ukazovateľ		CY	CZ	EE	HU	LT	LV	MT	PL	SI	SK	EU-25	EU-15	US	2005	2010	
Miera zamestnanosti (v %)	2003	69	65	63	57	61	62	54	51	63	58	63	64	71	67	70	

V miere zamestnanosti lisabonská stratégia venuje pozornosť aj znevýhodneným skupinám a rodovej rovnosti. V miere zamestnanosti žien sa pripúšťa a očakáva nižšia miera, avšak nemala by klesnúť pod 60 %. Priemer EU25 je 55 %, USA 66 % a na Slovensku je miera zamestnanosti 52 % (Tabuľka 4). Cieľovú hodnotu už prekračuje Rakúsko, Dánsko, Fínsko, Holandsko, Švédsko a dosahuje ju Cyprus. Horšia situácia ako na Slovensku je v Španielsku, Grécku, Taliansku, Maďarsku, Malte a v Poľsku. V miere zamestnanosti žien je postavenie Slovenska na prvý pohľad lepšie ako v miere zamestnanosti celkom. Zvyšovanie miery zamestnanosti žien je však do značnej miery závislé od cieľavedomej hospodárskej politiky v oblasti zosúladovania osobného a pracovného života žien, kde sa napríklad v oblasti predškolských zariadení situácia na Slovensku v posledných rokoch skôr zhoršuje ako zlepšuje.

Tabuľka 4: Miera zamestnanosti žien (15 až 65 rokov)

		EU - 15														
Ukazovateľ		AT	BE	DE	DK	ES	FI	FR	EL	IE	IT	LU	NL	PT	SE	UK
Miera zamestnanosti žien (v %)	2003	63	52	59	71	46	66	57	44	56	43	52	66	61	72	65

		EU - 10														Cieľ	
Ukazovateľ		CY	CZ	EE	HU	LT	LV	MT	PL	SI	SK	EU-25	EU-15	US		2005	2010
Miera zamestnanosti žien (v %)	2003	60	56	59	51	58	58	34	46	58	52	55	56	66		57	60

Z dôvodu starnutia obyvateľstva EÚ je ohrozená udržateľnosť dôchodkových a zdravotných systémov členských krajín a preto je ekonomicky odôvodnená snaha o oddiaľovanie odchodu do dôchodku a zavádzania modeli aktívneho starnutia. prejavilo sa to aj na vytýčení cieľa lisabonskej stratégie v miere zamestnanosti starších pracovníkov na relatívne nízkej úrovni 50 %. Ide o relatívne nový jav ako aj zavádzané riešenie a možno aj preto je situácia členských krajín veľmi rozdielna³. Najlepšia je vo Švédsku, kde už teraz pracujú siedmi z desiatich obyvateľov vo veku 55 až 65 rokov. Priemer EU25 je 40 %, USA 60 % a na Slovensku je táto miera zamestnanosti 25 % (Tabuľka 5). Cieľovú hodnotu okrem spomenutého Švédska prekračuje aj Dánsko, Veľká Británia a Estónsko. Horšia situácia ako na Slovensku je len v Slovinsku. Vzhľadom na vysokú našu mieru nezamestnanosti nemožno v strednodobom horizonte očakávať zdvojnásobenie miery zamestnanosti starších pracovníkov samotným pôsobením trhových síl.

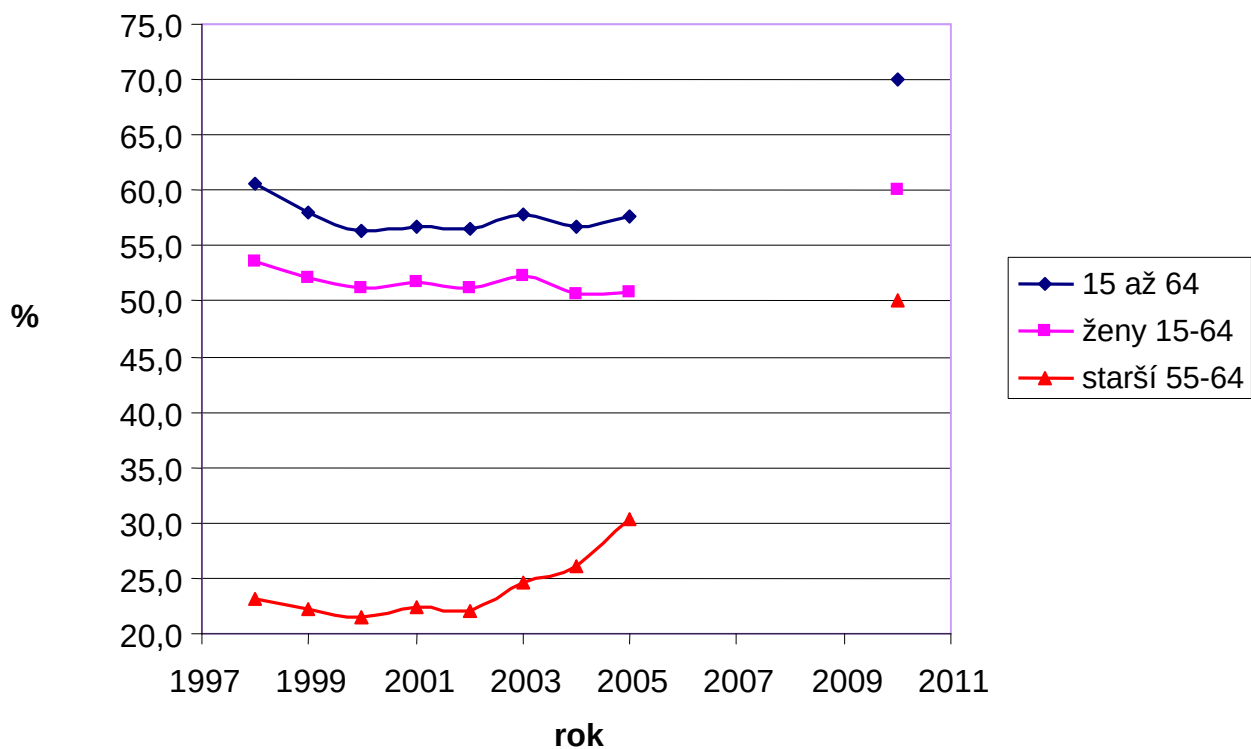
Tabuľka 5: Miera zamestnanosti starších pracovníkov (55 až 64 rokov veku)

		EU - 15														
Ukazovateľ		AT	BE	DE	DK	ES	FI	FR	EL	IE	IT	LU	NL	PT	SE	UK
Miera zamestnanosti starších pracovníkov (v %)	2003	30	28	40	60	41	50	37	42	49	30	30	45	51	69	56

		EU - 10														Cieľ	
Ukazovateľ		CY	CZ	EE	HU	LT	LV	MT	PL	SI	SK	EU-25	EU-15	US		2005	2010
Miera zamestnanosti starších pracovníkov (v %)	2003	50	42	52	29	45	44	33	27	24	25	40	42	60			50

O reálnosti splnenia našich cieľov miery zamestnanosti veľa napovedá dynamika týchto ukazovateľov v ostatných rokoch (Tabuľka 6 a Graf 1). V miere zamestnanosti starších pracovníkov nastalo v posledných rokoch zlepšenie, ktoré je spôsobené postupným oddiaľovaním odchodu do dôchodku a nie prechodom k cieľavedomou politike aktívnej staroby. Preto sa do značnej miery deje na úkor zamestnanosti žien a iných znevýhodnených skupín uchádzačov o zamestnanie. V miere zamestnanosti žien nastáva v posledných rokoch zhoršovanie situácie a miera zamestnanosti celkom má tendenciu k stagnácii.

³ Rozdielne sa chápe podstata zamestnanosti starších pracovníkov. Ozývajú sa aj názory, že práca starších pracovníkov je negatívny sociálny jav. My sa sústreďujeme na ekonomickú stránku tohto javu z pohľadu využitia ekonomického potenciálu krajiny a priaznivého vplyvu na verejné financie.

Graf 1: Miera zamestnanosti v SR**Tabuľka 6:** Miera zamestnanosti, SR v %

													Cieľ
Rok	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
15 až 64	60,6	58,0	56,3	56,7	56,5	57,9	56,7	57,7	-	-	-	-	70,0
ženy 15-64	53,6	52,1	51,1	51,8	21,2	52,3	50,6	50,9	-	-	-	-	60,0
starší 55-64	23,2	22,2	21,5	22,4	22,1	24,6	26,0	30,3	-	-	-	-	50,0

Zaujímavá je otázka zlučiteľnosti a náročnosti cieľov diskutovaných troch mier zamestnanosti v Slovenských podmienkach vysokej miery dlhodobej nezamestnanosti, dosiahnutého odborného vzdelania, snahy o zvýšenie natality a podielu populácie s vysokoškolským vzdelaním. V tabuľke 7 sú uvedené aktuálne miery zamestnanosti mužov, žien a populácie celkom vo vekovej štruktúre 15 až 19 rokov, 20 až 24, 25 až 54 a 55 až 65 rokov. Vzhľadom na potrebu udržania vysokého podielu populácie s ukončeným odborným vzdelaním nemožno očakávať zvýšenie ich miery zamestnanosti. Rovnako cieľ zvýšenia podielu vysokoškolákov a zvýšenia natality neumožňuje predpokladať zásadné zvýšenie miery zamestnanosti vekovej skupiny 20 až 24 rokov.

Tabuľka 7: Miery zamestnanosti mužov, žien a populácie, podľa vekových skupín v %

	Muži	Ženy	Muži a ženy spolu
15 – 19 rokov	3,3	2,8	3,1
20 – 24 rokov	52,2	42,7	47,6
25 – 54 rokov	80,9	68,9	74,9
55 – 65 rokov	47,1	16,0	30,3
Vekové skupiny spolu	64,6	50,9	57,7

V tabuľke 8 sú výsledky jednoduchého modelového prepočtu potrebnej miery zamestnanosti mužov, žien a populácie vo vekových skupinách za predpokladu dosiahnutia miery zamestnanosti starších pracovníkov 50 %, žien 60 % a celkovej 70 %. Dosiahnutie miery zamestnanosti mužov vo veku 25 až 54 rokov v roku 2010 na úrovni 94,2 až 100,0 % je v podmienkach vysokej dlhodobej nezamestnanosti veľmi ambicióznym až nereálnym cieľom a predovšetkým by plnenie tohto cieľa znamenalo dosiahnutie miery zamestnanosti u mužskej populácie vo vekovej skupine 55 až 65 rokov rovnako 72,2 až 100,0 %. Jedným východiskom môže byť zvyšovanie zamestnanosti v populácii vo veku 15 až 19 resp. 20 až 24 rokov, toto by však bolo v rozpore s už spomínaným cieľom zvyšovania vzdelania mladých ľudí a natality. Taktiež zvyšovanie zamestnanosti starších pracovníkov môže byť problematické, aj napriek ich ochote pracovať na skrátený úväzok. Je otázne, či ekonomika bude takéto pracovné miesta vytvárať a či budú obsadzované práve staršou pracovnou silou. Je veľmi pravdepodobné, že takéto miesta budú obsadzované práve pracovníkmi vo vekových kategóriách 15 až 24 rokov, čo by síce znamenalo, že miera zamestnanosti mužov vo vekovej kategórii 25 až 54 rokov by nemusela byť natoľko vysoká, ale zároveň by to pravdepodobne viedlo k nesplneniu cieľa miery zamestnanosti starších pracujúcich. Ďalším faktom ostáva, že ekonomická aktivita na trhu práce v roku 2005 dosiahla 2 645,7 tis. pracujúcich a nezamestnaných. Ak by mal byť splnený cieľ celkovej miery zamestnanosti na 70 %, znamenalo by to, že v roku 2010 by malo byť celkovo pracujúcich 2 703,6 tis., teda všetci súčasný pracujúci, všetci nezamestnaní a ešte naviac 57,9 tis. obyvateľov podľa nášho prepočtu. Pri prijatí predpokladu, že problémy s vysokou mierou dlhodobo nezamestnaných sa podarí vyriešiť a miera nezamestnanosti sa zníži na 3 až 5 percent, znamenalo by to, že celková ekonomická aktivita obyvateľov by mala vzrásť o približne 140 až 200 tis. obyvateľov, resp. miera ekonomickej aktivity by mala vzrásť o 3,5 až 5,5 percent.

Tabuľka 8: Miery zamestnanosti mužov, žien a populácie, podľa vekových skupín v %, prepočet na cieľ zamestnanosti v roku 2010

	Muži	Ženy	Muži a ženy spolu
15 – 19 rokov	3,3	2,8	3,1
20 – 24 rokov	52,2	42,7	47,6
25 – 54 rokov	94,2 – 100,0	79,6 – 85,5	89,8
55 – 65 rokov	72,2 – 100,0	7,4 – 31,1	50,0
Vekové skupiny spolu	80,1	60,0	70,0

Legenda: **Hrubým podčiarknuté** - cieľové miery zamestnanosti
Hrubým, kurzíva – variant 1 (100 % zam. u mužov vo vekovej kategórii 55 – 65 rokov)
Kurzíva – variant 2 (100 % zam. u mužov vo vekovej kategórii 25 – 54 rokov)

Záver

Dosiahnutie lisabonských cieľov miery zamestnanosti je v podmienkach Slovenska :

- významné pre trvalú udržateľnosť verejných financií,
- dôležitý faktor ekonomického potenciálu,
- nutná podmienka zásadného zníženia daňového a odvodového zaťaženia,
- ich dosiahnutie je na Slovensku nepravdepodobné.

Pravdepodobnosť ich dosiahnutia sa zvýši:

- zvyšovaním zamestnanosti,
- skúmaním demografického starnutia s implikáciami,
- celoživotným vzdelávaním,
- využívaním flexibilných foriem zamestnávania,
- investovanie do vysoko kvalifikovanej pracovnej sily,
- nájdením rovnováhy medzi flexibilitou a istotou,
- zosúladienie osobného a pracovného života,
- aktívnym starnutím.

Literatúra:

Prijat' výzvy – Lisabonská stratégia rastu a zamestnanosti. Správa pracovnej skupiny na vysokej úrovni pod predsedníctvom Wima Koka, Európska komisia, november 2004, 51 s.
http://ec.europa.eu/growthandjobs/pdf/kok_report_sk.pdf

Prílohy:

Na základe Lisabonskej stratégie sa majú vykazovať nasledovné ukazovateľ v rámci správ (napr. Wim Kok report) a to počas nasledujúcich 3 rokov (2004 - 2007). 15 základných ukazovateľov:

1. **GDP per capita in PPS:** HDP na obyvateľa v parite kúpnej sily: Štrukturálny ukazovateľ Eurostat-u, skupina General Economic background (ecobac)
 - a. **Ukazovateľ/Dáta:** Hodnoty hrubého domáceho produktu sú v parite kúpnej sily na obyvateľa. Dáta sú poskytnuté členskými štátmi podľa normy ESA 95, ročné dáta, národné účty a hlavné agregáty, časový rad je vykazovaný od roku 1995, pre niektoré krajiny od roku 1991
 - b. **Pravidlá štatistického spracovania:** Národné účty sú spracované v súlade so európskym systémom účtov (ESA 95), prijatým vo forme regulácie Rady 25. júna 1996 N° 2223/96. Regulácie

štatistickej dôveryhodnosti: regulácia Rady (CE) N° 322/97 zo 17 februára 1997 (OJ No L 52/1) a regulácia Rady (EURATOM, EEC) N° 1588/90 z 11 júna 1990 (OJ No L 151/ 1)

- c. **Metodológia:** Ukazovateľ je vyjadrený ako relácia k EU25 = 100, resp EU15 = 100 teda ako index. Výpočet HDP v parite kúpnej sily na obyvateľa je vytvorený na porovnanie ekonomickej aktivity v jednotlivých krajinách rôznej veľkosti. Nieje vhodný na porovnávanie v čase. Vzhľadom na to že niektoré dáta členských krajín niesu k dispozícii tak sú urobené odhady ich hodnôt.

d.

2. **Labour productivity per person employed:** Produktivita práce na zamestnaného: Štrukturálny ukazovateľ Eurostat-u, skupina General Economic background (ecobac)

- a. **Ukazovateľ/Dáta:** Ide o porovnanie rastu jednotkových nákladov práce k odmenám zamestnancov a produktivity práce tak aby zobrazili aký stupeň odmien zamestnancov je v relácii k ich produktivite práce. Teda ide o zobrazenie relácie koľko je zamestnanec platený ku produktivite jeho práce. Dáta sú poskytnuté členskými krajinami podľa normy ESA 95, ročné dáta, národné účty a hlavné agregáty, časový rad je vykazovaný od roku 1995, pre niektoré krajiny od roku 1991
- b. **Pravidlá štatistického spracovania:** Národné účty sú spracované v súlade so európskym systémom účtov (ESA 95), prijatým vo forme regulácie Rady 25. júna 1996 N° 2223/96. Regulácie týkajúce štatistickej dôveryhodnosti: regulácia Rady (CE) N° 322/97 zo 17 februára 1997 (OJ No L 52/1) a regulácia Rady (EURATOM, EEC) N° 1588/90 z 11 júna 1990 (OJ No L 151/ 1)
- c. **Metodológia:** Na výpočet produktivity práce sa používa HDP (ESA95, 8.89) v bežných cenách meraný v parite kúpnej sily a zahrňuje dane a dotácie na produkciu a import. Počet zamestnaných osôb je rovný úplnej zamestnanosti (podľa konceptu národných účtov ESA95, 11.11) zahrňujúcej všetky osoby, zamestnancov aj samo-zamestnaných, zapojené v produktívnej aktivite ktorá zapadá do rámca systému účtov. Za zamestnancov sa považujú: zamestnanci podľa zamestnaneckej zmluvy, štátni zamestnanci (verejný zamestnanci), personál ozbrojených síl, majitelia korporácií a podnikov ak v nich pracujú... (pozri ESA95 11.12). Osoby dočasne neprítomné v práci z dôvodu choroby, zranenia, dovolenky, prázdnin, štrajku či vzdelávania sú takisto považované za zamestnancov. Ukazovatele zamestnanosti sa môžu líšiť od štatistiky pracovných síl kvôli rôznemu konceptu ich vykazovania. Ukazovateľ je vyjadrený ako relácia k EU25 = 100, resp EU15 = 100 teda ako index.

3. **Employment rate (%):** Miera zamestnanosti (%):

4. **Employment rate females (%):** Miera zamestnanosti žien(%):

5. **Employment rate of older workers (%):** Miera zamestnanosti starších pracovníkov(%): Štrukturálny ukazovateľ Eurostat-u, skupina Labour market

- a. **Ukazovateľ/Dáta:** Zamestnané osoby vo veku od 15-64 rokov ako podiel na celkovej populácii v rovnakej vekovej skupine - harmonizované s dátami obyvateľstva a národnými účtami. Pre starších pracovníkov je vekový interval 55-64 rokov. EU25 od roku 1997. Dáta sú vykazované 12 týždňov po skončení príslušného roka ako ročný priemer.
- b. **Pravidlá štatistického spracovania:** Prieskum pracovných síl Európskej Únie je riadený legislatívnym aktom európskej rady a parlamentu ako aj komisiou pre implementáciu. Základnou legislatívou je regulácia Rady (CE) N° 577/98 z 9. marca 1998 na organizáciu prieskumu pracovných síl v EU (OJ No L 77/3). Regulácie týkajúce štatistickej dôveryhodnosti: Regulácia rady (CE) N° 322/97 z 17. februára 1997 (OJ No L 52/1) a Regulácia rady (EURATOM, EEC) N° 1588/90 z 11. Júna 1990 (OJ No L 151/ 1)
- c. **Metodológia:** Zamestnané osoby sú všetky osoby vo veku 15 rokov a viac ktoré pracovali za mzdu alebo zisk v danom týždni. Miera zamestnanosti osôb vo veku 15 až 64 rokov resp. 55 až 64 rokov je podiel zamestnaných osôb v tomto vekovom intervale ku celkovej populácii v tomto intervale (15-64 resp. 55-64). Ročné dáta sú odvodené zo štvrtročných prieskumov európskych pracovných síl. Agregáty EU-zóny sú odvodené z dát získaných na národnej úrovni. Keďže všetky dáta sú vyjadrené v absolútnych číslach (v počte osôb) niesu použité žiadne váhy, agregátne počty sú rátané sčítaním všetkých národných dát. Miery zamestnanosti sú odvodené na základe ročných priemerov populácie(zamestnancov). Ročné priemery sú rátané na základe štvrtročných dát ako jednoduché priemery. Štvrtročné prieskumy pracovných síl sú použité keď sú k dispozícii, inak sú použité interpolácie medzi dvoma najbližšími štvrtrokmi pre chýbajúci štvrtrok. Interpolácie niesu zverejnené ale sa používajú pre ročné priemery, ak chýbajú dáta pre danú krajinu a štvrtrok, na výpočet EU agregátov.

6. **Educational attainment (20-24)(%)**: Dosiahnutý stupeň vzdelanosti mladých ľudí vo veku 20-24 rokov: Štrukturálny ukazovateľ Eurostat-u, skupina Labour market
- a. Ukazovateľ/Dáta: Osoby vo veku 20 až 24 rokov ktoré dosiahli stredoškolské vzdelanie ako percento populácie v rovnakej vekovej skupine. Časový rad je publikovaný od roku 1992 pre členské krajiny, a od 1995 postupne aj pre krajiny prístupujúce.
 - b. Pravidlá štatistického spracovania: Prieskum pracovných síl Európskej Únie je riadený legislatívnym aktom európskej rady a parlamentu ako aj komisiou pre implementáciu. Základnou legislatívou je regulácia Rady (CE) N° 577/98 z 9. marca 1998 na organizáciu prieskumu pracovných síl v EU (OJ No L 77/3). Regulácie týkajúce štatistickej dôveryhodnosti: Regulácia rady (CE) N° 322/97 z 17. februára 1997 (OJ No L 52/1) a Regulácia rady (EURATOM, EEC) N° 1588/90 z 11. júna 1990 (OJ No L 151/ 1)
 - c. Metodológia: Osoby vo veku 20 až 24 rokov ktoré dosiahli stredoškolské vzdelanie ako percento populácie v rovnakej vekovej skupine, min level ISCED 3 resp. 4. ISCED 3 – Ide o stredoškolský stupeň vzdelania, začínajúci po 9 ročnej základnej školskej dochádzky a je typický pre systémy s povinnou školskou dochádzkou. Typický vek vstupu študentov do tohto stupňa vzdelávania je 15 rokov. Na tento stupeň nadväzuje, ak je vzdelávanie ukončené maturitou, stupeň ISCED 5 (bakalárske štúdium). ISCED 3A, 3B, 3C sú stupne vzdelania (u nás ukončené maturitnou skúškou) kde je predpoklad na pokračovanie vo vzdelávaní na stupni ISCED 5A resp 5B teda plnom vysokoškolskom štúdiu s možnosťou pokračovať na úroveň ISCED 6 a získania vedeckej hodnosti. Stupeň ISCED 4 je post-sekundárne štúdium (po ukončení strednej školy) ale nejde o terciálne (vysokoškolské) štúdium. Spadá sem rekvalifikačné štúdium, kurzy a iné štúdium, ktoré však nezvýši kvalifikáciu študenta nad stupeň ISCED 3, 3A, 3B, 3C. Vykazované sú osoby trvale žijúce v krajine, vojaci konajúci základnú vojenskú službu a osoby žijúce v hromadných ubytovacích zariadeniach (liečebných ústavoch a i.) sú vylúčené.
7. **Research and development expenditure (% of GDP)**: Výdaje na výskum a rozvoj ako percentuálny podiel na HDP: Štrukturálny ukazovateľ Eurostat-u, skupina Innovation and Research
- a. Ukazovateľ/Dáta: Výdaje na výskum a rozvoj sa skladajú zo: súkromných výdavkov na výskum a rozvoj, vysokoškolské vzdelávanie vo výskume a rozvoji, vládne výdaje a výdaje súkromných neziskových spoločností.
 - b. Pravidlá štatistického spracovania: Do roku 2003 boli dáta zhromažďované na základe gentlemankej dohody od roku 2003 sú zhromažďované na základe regulácie N° 753/2004 (OJ L 118, strana 23, 23 Apríl 2004). Regulácie týkajúce štatistickej dôveryhodnosti: Regulácia rady (CE) N° 322/97 z 17. februára 1997 (OJ No L 52/1) a Regulácia rady (EURATOM, EEC) N° 1588/90 z 11. júna 1990 (OJ No L 151/ 1)
 - c. Metodológia: Ukazovateľ je rátaný na základe súčasnej hodnoty ECU/EUR. Dáta sú získavané na základe ročných dotazníkov eurostatu. Za štatistické jednotky sa považujú: Vládny sektor (inštitúcie, štátne orgány..), Sektor vzdelávania (univerzity, vzdelávacie centrá...), neziskové súkromné organizácie a súkromné spoločnosti. Dáta sú zasielané z členských krajín eurostatu. EU a iné agregáty sú rátané ako suma národných dát podľa sektorov. V prípade že dáta chýbajú eurostat urobí odhad pre daný čas a premennú. Vzhľadom na povahu tohoto ukazovateľa (iná štruktúra hospodárstva členských krajín EU, či vykazovanie výdavkov inštitúcii o spoločností) treba brať na zreteľ značné skreslenie vykazovaných čísiel.
8. **Business investment**: Investície: Štrukturálny ukazovateľ, skupina Economic Reform
- a. Ukazovateľ/Dáta: Ukazovateľom sa vykazujú investície (formovanie hrubého fixného kapitálu) súkromného sektora ako percentuálny podiel na HDP.
 - b. Pravidlá štatistického spracovania: Národné účty sú spracované v súlade so európskym systémom účtov (ESA 95), prijatým vo forme regulácie Rady 25. júna 1996 N° 2223/96. Regulácie štatistickej dôveryhodnosti: regulácia Rady (CE) N° 322/97 zo 17 februára 1997 (OJ No L 52/1) a regulácia Rady (EURATOM, EEC) N° 1588/90 z 11 júna 1990 (OJ No L 151/ 1)
 - c. Metodológia: Formovanie hrubého fixného kapitálu je definované (GFCF, ESA 1995, 3.102) ako zisk domácich výrobcov znížený o odpisy fixného kapitálu v danom období a zvýšený o nárast hodnoty neprodukovaných aktív produkciou výrobcu resp. inštitucionálnej jednotky. Odpisy sú brané ako záporný zisk. Súkromný sektor sa skladá zo súkromných nefinančných a finančných korporácií, domácností a neziskových organizácií slúžiacich domácnostiam teda všetky sektory okrem vládneho sektora (ESA 1995, 2.21 - 2.88). Pre výpočet podielu hrubého fixného kapitálu na HDP sa používajú bežné ceny prepočítané v Euro (pred 1999 v ECU).

9. Comparative price levels: Porovnanie cenovej úrovne, Štrukturálny ukazovateľ, skupina Economic reform

- a. Ukazovateľ/Dáta: Parity kúpnej sily sú rátane Eurostat-om, kde Eurostat je zodpovedný za koordináciu a kalkuláciu finálnych parít na základe dát a cenových prieskumov Národných Štatistických Inštitútov (NSIs).
- b. Pravidlá štatistického spracovania: Aplikovaná metodológia a publikovaná úroveň agregácie je dohodnutá na základe dohody PPP Working Party (WP). Vypočítané hodnoty PPP eurostat-om sú zasielané delegátom WP na verifikovanie predtým než sú publikované. Legislatíva únie na spracovanie PPP je v procese prípravy. Regulácie štatistickej dôveryhodnosti: regulácia Rady (CE) N° 322/97 zo 17 februára 1997 (OJ No L 52/1) a regulácia Rady (EURATOM, EEC) N° 1588/90 z 11 júna 1990 (OJ No L 151/ 1)
- c. Metodológia: Ukazovateľ je rátaný na základe porovnania Parity kúpnej sily obyvateľstva na základe výdavkov na spotrebu, a vyrátaný index ($EU15/25 = 100$) slúži na porovnanie cenovej konvergencie jednotlivých členských štátov únie. PPS sú rátané na základe spotrebného koša ktorý zodpovedá nákupnému správaniu a spotrebe tovarov a služieb v jednotlivých krajinách únie. Jednoduché cenové relácie na úrovni jednotlivých produktov sú agregované na základe váženého priemeru do cenových relácií na úrovni skupín produktov s postupnou agregáciou na finálnu spotrebu a v poslednej rade na HDP. Na opravu chýb v procedúre numerického výpočtu PPPs sa obvykle volí jedna krajina ako báza a jej PPS je rovné jednej. V komparácii OECD sú Spojené Štáty Americké vždy báza. Pre strednú a východnú Európu sa v minulosti používalo Rakúsko ako bázička krajina. Algebra a metóda použitá eurostat-om zaručuje že bázička krajina je invariantom, čo znamená že cenové relácie medzi krajinami sú nemenné nech je zvolená ktorákoľvek krajina za bázu. Každopádne v kontexte EU nieje vhodné zvoliť niektorú krajinu/menu za bázičku krajinu/menu pre výpočet PPPs. Preto PPPs sú škálované tak že EU-HDP v PPS je rovné EU-HDP v eure. To znamená že pre priemer EU sa 1 PPS rovná 1 Euro. Ukazovatele PPPs a súvisiace ekonomické ukazovatele sú konštruované na zobrazenie priestorovej komparácie a nie na komparáciu v čase. Preto každé porovnanie v čase týchto ukazovateľov musia na toto brať zreteľ.

10. At-risk-of-poverty rate: Miera risku chudoby, Štrukturálny ukazovateľ, skupina Social Cohesion (socohe)

- a. Ukazovateľ/Dáta: Ukazovateľ je definovaný ako percento obyvateľstva s disponibilným príjmom pod prahom chudoby, ktorý je stanovený ako 60% priemerného disponibilného príjmu po sociálnych transferoch. Tento podiel je rátaný pred sociálnymi transfermi (príjem aj s dôchodkami ale bez iných sociálnych transferoch) a so sociálnymi transfermi (celkový príjem)
- b. Pravidlá štatistického spracovania: Regulácie štatistickej dôveryhodnosti: regulácia Rady (CE) N° 322/97 zo 17 februára 1997 (OJ No L 52/1) a regulácia Rady (EURATOM, EEC) N° 1588/90 z 11 júna 1990 (OJ No L 151/ 1). Členovia EU spolupracujúci na 'European Community Household Panel survey' na základe gentlemen'skej dohody sa uzniesli posilniť kontext 'otvorených metód koordinácie' na Lisabonskom zhromaždení.
- c. Metodológia: Kmeňové dáta pochádzajú na základe 'European Community Household Panel (ECHP)', užívateľská databáza, ktorá obsahuje priebežné mikro-dáta o domácnostiach a osobách. Tieto panelové dáta sú získavané na základe prieskumu reprezentatívnych domácností a jednotlivcov. Vzhľadom nato že ide o výberovú vzorku vyrátaný podiel domácností ktoré sú v riziku chudoby je ovplyvnený veľkosťou výberového súboru. Vzhľadom nato že je legislatíva o štatistickom vykazovaní 'Príjmu a Životných podmienok (EU-SILC)' v stave pripravovania, na zaručenie maximálnej konzistencie dát eurostat používa dáta aj z národných zdrojov členských krajín.

11. Long-term unemployment rate: Dlhodobá miera nezamestnanosti, Štrukturálny ukazovateľ, skupina Social Cohesion (socohe)

- a. Ukazovateľ/Dáta: Ukazovateľ dlhodobej nezamestnanosti je rátaný ako podiel dlhodobo nezamestnaných (12 mesiacov a viac) na ekonomicky aktívnom obyvateľstve – harmonizované s národnými mesačnými odhadmi nezamestnanosti.
- b. Pravidlá štatistického spracovania: Základnou legislatívou je regulácia Rady ([CE](#)) N° 577/98 z 9. marca 1998 na organizáciu prieskumu pracovných síl v EU (OJ No L 77/3). Regulácie štatistickej dôveryhodnosti: regulácia Rady (CE) N° 322/97 zo 17 februára 1997 (OJ No L 52/1) a regulácia Rady (EURATOM, EEC) N° 1588/90 z 11 júna 1990 (OJ No L 151/ 1)

- c. **Metodológia:** Za nezamestnané osoby sa považujú všetky osoby vo veku 15 -74 rokov ktoré neboli zamestnané v referenčnom týždni, aktívne si hľadali prácu predchádzajúce 4 týždne a sú okamžite pripravený začať pracovať resp. do 2 týždňov. Dĺžka nezamestnanosti je definovaná ako dĺžka hľadania si práce, alebo ako dĺžka periódy odkedy bola osoba naposledy zamestnaná. Dlhodobá miera nezamestnanosti je podiel nezamestnaných osôb posledných 12 mesiacov a viac na celkovom počte aktívnych osôb na trhu práce. Actívne osoby sú tie ktoré sú zamestnané aj nezamestnané (ekonomicky aktívne obyvateľstvo). Agregáty pre Eu-zónu (EU15, EU25...) sú odvodené od dát získaných na národnej úrovni. Miera dlhodobej nezamestnanosti je odvodená od ročných dát prieskumu obyvateľstva. Ročné údaje sú odvodené od štvrtročných dát prieskumu obyvateľstva.
12. **Dispersion of regional employment rate:** Regionálny rozptyl miery nezamestnanosti – chyba metodológia na webe.
13. **Greenhouse gas emissions (index base year = 100):** Emisie skleníkových plynov (index základný rok = 100), Štrukturálny ukazovateľ, skupina Environment
- Ukazovateľ/Dáta:** Ide o následovné plyny a zlúčeniny: CO₂, N₂O, CH₄, HFC_s, PFC_s, SF₆. Zber dát a ich predloženie výboru je riadené rozhodnutím N° 280/2004/EC Európskeho parlamentu a rady z 11 februára 2004.
 - Pravidlá štatistického spracovania:** Legálnym základom pravidiel štatistického spracovania je rozhodnutie N° 280/2004/EC a N° 358/2002/EC so súhlasom UNFCCC a Kyótskym protokolom, 1997. Regulácie štatistickej dôveryhodnosti: regulácia Rady (CE) N° 322/97 zo 17 februára 1997 (OJ No L 52/1) a regulácia Rady (EURATOM, EEC) N° 1588/90 z 11 júna 1990 (OJ No L 151/ 1)
 - Metodológia:** Ukazovateľ je rátný na základe Kyótskeho koša (Kyoto basket) šiestich skleníkových plynov vážených na základe prispievaniu ku tvorbe globálneho otepľovania, v 1000 ton CO₂ ekvivalentu, delené emisiou v základnom roku. Ide o následovné plyny a zlúčeniny: CO₂, N₂O, CH₄, HFC_s, PFC_s, SF₆. Základné údaje sú ročné odhady emisií skleníkových plynov oznamované na základe revidovaného medzivládneho panelu klimatických zmien ([IPCC](#)), rovnako ako aj európskej environmentálnej agentúry. Agentúra spracováva ročné údaje z 25 členských štátov, ktoré štáty vykazujú sekretariátu Spojených národov, Rámcová konvencia klimatických zmien ([UNFCCC](#)).
14. **Energy intensity of the economy:** Celková vnútroštátna spotreba energie ako podiel na HDP, Štrukturálny ukazovateľ, skupina Environment
- Ukazovateľ/Dáta:** Tento ukazovateľ je rátný ako podiel vnútroštátnej spotreby energie na HDP. HDP je vykazované na základe ESA95. Hrubá vnútorná spotreba energie je meraná v ekvivalentoch kilogramu oleja (kgoe) a HDP v 1000 Euro, výsledný pomer je kgoe na 1000 EUR. HDP je rátný v stálych cenách.
 - Pravidlá štatistického spracovania:** Poskytovanie dát je na základe gentlemanskej dohody medzi členmi EU, kandidátskymi krajinami a krajinami EFTA. Regulácie štatistickej dôveryhodnosti: regulácia Rady (CE) N° 322/97 zo 17 februára 1997 (OJ No L 52/1) a regulácia Rady (EURATOM, EEC) N° 1588/90 z 11 júna 1990 (OJ No L 151/ 1)
 - Metodológia:** Vzhľadom na to že údaje sú pravidelne vykazované medziročne málokedy dochádza k odhadom. Údaje sú vykazované spoločnými dotazníkmi Eurostat-u a Medzinárodnej energetickej agentúry (patrí OECD). To znamená že metodológia EU je kompletne harmonizovaná s krajinami OECD, vrátane USA a Japonska. Údaje sú vykazované 14 mesiacov po ukončení referenčného roka. HDP je rátný v stálych cenách.
15. **Volume of transport:** Množstvo nákladnej prepravy pomerne ku HDP, Štrukturálny ukazovateľ, skupina Environment
- Ukazovateľ/Dáta:** Index množstva vnútornej nákladnej prepravy ako podiel ku HDP. Meraný v tonokilometroch/HDP (v stálych cenách z roku 1995 a v Eurách), 1995 = 100.
 - Pravidlá štatistického spracovania:** regulácia (EC) N° 1172/98, direktíva 80/1177/EEC nahradená reguláciou (EC) N° 91/2003, direktíva 80/1119/EEC, regulácia (EC) N° 91/2003. Pokiaľ údaje nebolo možné získať na základe regulácií a direktív, sú použité údaje národných štatistických inštitúcií, ECMT, UNECE, UIC, DG pre Energie a prepravu. Regulácie štatistickej dôveryhodnosti: regulácia Rady (CE) N° 322/97 zo 17 februára 1997 (OJ No L 52/1) a regulácia Rady (EURATOM, EEC) N° 1588/90 z 11 júna 1990 (OJ No L 151/ 1)

- c. Metodológia: Za vnútornú prepravu sa považuje železničná, lodná a pozemná preprava. Železničná a lodná preprava sa ráta na prejazd v rámci krajiny bez ohľadu na národnú príslušnosť vozidla či plavidla. Cestná preprava sa ráta na základe pohybu registrovaných vozidiel v danej krajine.

Pre ďalšie podrobnejšie informácie viz:

http://epp.eurostat.ec.eu.int/portal/page?_pageid=1133,1418053,1133_1418058&_dad=portal&_schema=PORTAL

Konštrukcia modelov HERMIN a CGE pre hodnotenie Národného Strategického Referenčného Rámca (NSRR)

Vladimír Kvetan, Martin Mlýnek, Marek Radvanský, Viliam Páleník ¹

Abstract: Planning policy evaluation and verification of assumed policy impact on economic development is main purpose of CGE and HERMIN model construction. Preparation of National Development Plan (NDP) of Slovak republic (NSRR), co financed by European Union funds, was main cause for construction of these models of Slovak economy. In the first part of this paper construction of HERMIN model is discussed, in second part construction of CGE model and in last part evaluation of three variants of Slovaks NDP's. As previous studies shown, main question of NDP is not magnitude of reallocations but its structure. In this respect, the results of this paper seem to be consistent with other empirical findings.

Key words: computable general equilibrium model, CGE, econometric model, HERMIN, economic development, planning policy evaluation.

V tomto príspevku sa budeme venovať použitiu štruktúrneho modelu HERMIN a CGE modelu hodnotení NSRR. V prvej časti článku je stručne popísaná s konštrukcia modelu typu HERMIN. V ďalšej časti je charakterizovaná konštrukcia CGE modelov. V poslednej časti bude popísané hodnotenie NSRR. Ťažisko tejto časti je v popise vnútorných modelových mechanizmov vyhodnocovania. V tejto časti budú hodnotené tri varianty NSRR tak ako boli postupne vypracovávané zverejňované Partnerstvom pre NSRR .

Konštrukcia modelu HERMIN

Modely typu HERMIN sú vyvíjané od osemdesiatych rokov dvadsiateho storočia a vychádzajú z modelovej štruktúry modelu HERMES, ktorý bol vyvinutý európskou komisiou začiatkom osemdesiatych rokov (d'Alcantara and Italianer, (1982)). Model HERMIN predstavuje zjednodušenú verziu modelu HERMES tak, aby bol aplikovateľný v malých otvorených ekonomikách akou ekonomika Slovenska nepochybne je.. Zjednodušenie modelu podstatne znižuje jeho náročnosť na dátovú základňu, pričom z veľkej

¹ Autori sú pracovníkmi Ekonomického ústavu SAV

miery zachováva mechanizmy a evaluačné schopnosti východiskového modelu HERMES. Tento model bol aplikovaný v Grécku, Írsku, Portugalsku, Severnom Írsku a Španielsku a v rámci prístupových rokov na evaluáciu využívania prístupových fondov v Českej Republike, Estónsku, Rumunsku a Slovinsku.

Model HERMIN je určený pre malé otvorené ekonomiky, zohľadňuje štruktúru nástrojov kohéznej ekonomiky a preto musí spĺňať tieto základné predpoklady:

- Ekonomika musí byť dezagregovaná na malý počet sektorov, ktoré umožňujú identifikovať kľúčové štrukturálne zmeny v ekonomike počas skúmaného obdobia²
- Musí mať vyšpecifikovaný mechanizmus pomocou ktorého je ekonomika prepojená s „vonkajším svetom“. Tento mechanizmus by mal zachytávať medzinárodnú výmenu tovarov a služieb, inflačné prenosy, migráciu pracovnej sily ako aj priame zahraničné investície. Vonkajšia (alebo svetová) ekonomika je veľmi dôležitým priamym aj nepriamym faktorom, ovplyvňujúcim ekonomický rast a konvergenciu krajiny k rozvinutým krajinám EU
- Produkcia jednotlivých hospodárskych sektorov zahrnutých v modeli je vyjadrená pomocou produkčných funkcií v špecifickom tvare (CES, Cobb-Douglas)
- Tvorca a používateľ modelu musí rozpoznať možný konflikt medzi súčasným stavom ekonomiky, ktorý na základe historických dát popisuje model HERMIN, a budúcim očakávaným stavom ekonomiky približujúcemu sa prostrediu dominovanému jednotným spoločným trhom EU³

Model HERMIN Slovenskej republiky, ISWE - HERMIN 05 jednoznačne spĺňa tieto predpoklady. Štruktúra modelu je prispôsobená tak, aby popisovala štruktúru ekonomiky Slovenskej republiky. Model je založený na popise troch základných blokov: bloku produkcie, bloku absorpcie a bloku rozdelenia príjmov. Produkčný sektor je v modeli dezagregovaný do štyroch sektorov: sektora priemyslu, trhových služieb, netrhových služieb, a pôdohospodárstva.

Sektor priemyslu pozostáva z agregátu sektorov ťažba nerastných surovín (C), priemyselná výroba (D) a stavebníctvo (F) podľa OKEČ. Sektor výroby a rozvodu elektriny, plynu a vody (E) je v rámci štruktúry modelu zaradený do sektoru trhových služieb. Hlavným

² Počet zvolených sektorov ekonomiky závisí od údajovej základne hodnotiteľa NDP

³ Predovšetkým musia byť zapracované nové aspekty ekonomiky (mobilita pracovnej sily, očakávané štrukturálne zmeny...) v hodnotení ex-ante, ktoré vyplývajú z kohéznej politiky a európskeho jednotného trhu, ak tieto informácie nie sú zachytené modelom HERMIN na základe historických dát

dôvodom tejto modifikácie je, že produkčná funkcia tohto sektoru vykazuje parametre produkčných funkcií pre sektor služieb. Ponechanie výroby a rozvodu elektriny, plynu a vody v sektore priemyslu by vnieslo významnú poruchu do agregovanej produkčnej funkcie pre celý sektor na základe analýzy, ktorú autori vykonali pri konštrukcii modelu⁴. Rovnica produkcie pre sektor priemyslu je stochastická s vysvetľujúcimi premennými zahraničný dopyt, domáci dopyt, jednotkové náklady práce, porovnanie domácej a zahraničnej cenovej hladiny a trendová zložka. Zásoba kapitálu je rátaná ako súčet investícií a zásoby kapitálu v minulom období, znížené o odpisy. Miera odpisov predstavuje konštantnú úroveň 2% ročne.

Sektor trhových služieb pozostáva zo sektorov: výroba a rozvod elektriny, plynu a vody (E), veľkoobchod a maloobchod, oprava motorových vozidiel, motocyklov a spotrebného tovaru (G), hotely a reštaurácie (H), doprava, skladovanie, pošty a telekomunikácie (I), finančné sprostredkovanie (J) a nehnuteľnosti, prenájom a obchodné činnosti (K). Produkcia v odvetví netrhových služieb je závislá od zahraničného a domáceho dopytu.

Sektor netrhových služieb pozostáva zo sektorov Verejná správa a obrana, povinné sociálne zabezpečenie (L), Školstvo (M), Zdravotníctvo a sociálna pomoc (N), Ostatné spoločenské, sociálne a osobné služby (O), Činnosti domácností (P) a Exteritoriálne organizácie a združenia (Q). Produkcia v sektore netrhových služieb je závislá od domáceho dopytu spolu s trendovou zložkou a nie je exogénnou (riadiacou) premennou modelu, tak ako to predpokladá špecifikácia základného modelu HERMIN. Dôvodom na popísanie produkcie v tomto sektore pomocou stochastickej rovnice bolo prijatie predpokladu závislosti produkcie od domáceho dopytu. Predovšetkým v sektore zdravotníctva a školstva sa na základe spustených a pripravovaných reforiem vytvára trhové prostredie.

Pôdohospodársky sektor pozostáva zo sektora poľnohospodárstvo, poľovníctvo a lesníctvo (A) a rybolov, chov rýb (B). Produkcia v sektore je závislá na domácom dopyte. Dopyt po práci a investíciách je vo všetkých sektoroch odvodený na základe nákladovej minimalizácie C-D produkčnej funkcie.

Blok absorpcie predstavuje dopytovú stránku ekonomiky. Konečná spotreba domácností je v modeli vyjadrená rovnicou závislosti spotreby od disponibilných príjmov domácností. Celkový domáci dopyt je súčtom konečnej spotreby domácností, spotreby verejnej

⁴ Analýza spočívala v odhade elasticít produkčných funkcií za jednotlivé sektory podľa OKEČ a následné zoskupenie sektorov na základe korelačnej analýzy do agregovaných sektorov (množín). Analýza poukázala na to že charakteristiky produkčnej funkcie sektora (E) výroby a rozvodu elektriny, plynu a vody sú v zhode so sektormi služieb (sektory G, H I, J, K podľa OKEČ)

správy, investícií a zmeny stavu zásob. Saldo zahraničného obchodu je riešené neštandardne pomocou identity ako rozdiel celkového výstupu ekonomiky a domáceho dopytu.

Blok rozdelenia príjmov v ekonomike je modelovaný prevažne identitami. Dôraz je kladený predovšetkým na ceny výdavkov verejného sektora, ktoré sú modelované stochastickou rovnicou, kde vysvetľujúce premenné sú ceny výstupov, dovozné ceny a miera nepriameho zdanenia. Disponibilný príjem domácností je vyjadrený ako príjem za ekonomiku znížený o priame dane a navýšený o transfery domácnostiam. Saldo bežného účtu je vyrátané pomocou identity ako súčet salda zahraničného obchodu a príjmov výrobných faktorov zo zahraničia. Deficit verejných financií vzniká rozdielom medzi príjmami a výdavkami verejného sektora. Verejný dlh sa rovná súčtu verejného deficitu a dlhu z predchádzajúceho obdobia.⁵

Konštrukcia CGE modelu

Pre hodnotenie NSRR bol skonštruovaný rekurzívno-dynamický CGE model. Údajovú základňu modelu predstavuje matica spoločenského účtovníctva (SAM) za rok 2000. Na rozdiel od ekonometrických modelov, CGE modely predstavujú deterministické modely a sú založené na teórii a predpokladoch neoklasickej ekonomickej školy.

Údajová základňa, SAM matica, je agregovaná na päť produkčných sektorov, a to sektor pôdohospodárstva (A a B), priemyslu (C, D, E), stavebníctva (F), trhových služieb (G, H, I, J, K) a netrhových služieb (L, M, N, O, P, Q). Predpokladá sa jedna agregátna domácnosť (HC), ktorej správanie je popísané Cobb-Douglasovou funkciou užitočnosti. Dopyt po investíciách (INV) je popísaný Lieontieffovou funkciou, rovnako ako spotreba verejnej správy (GC). Zahraničný obchod je popísaný pomocou Armingtonovej koncepcie, teda dovoz (IM) je spolu s domácou produkciou (DP) agregovaný CES (constant elasticity of substitution) funkciou na celkovú domácu ponuku (A) a celková produkcia (O) je transformovaná na vývoz (EX) a domácu ponuku (DP) na základe CET (constant elasticity of transformation) funkcie.

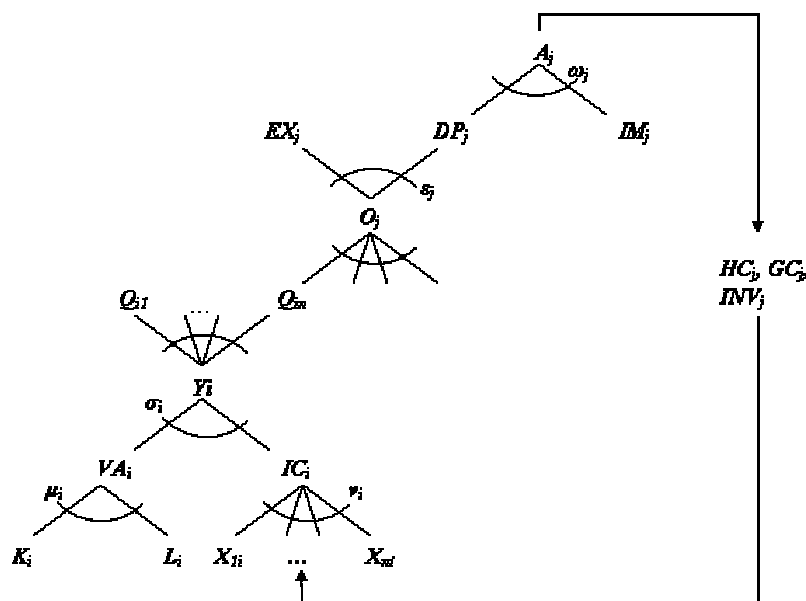
Produkcia (Y) je v modeli riešená na základe vnorenej produkčnej funkcie pre jednotlivé sektory. Funkcia má nasledovný tvar:

$$Y = \text{Leontieff}(\text{CES}(K, L), \text{Leontieff}(X_1, X_2, \dots, X_5)) \quad (1)$$

⁵ Pre bližšie informácie pozri: Ekonomický časopis SAV, 2006, č. 7, s.

Horná úroveň produkčnej funkcie (PF) predstavuje Leontieffovu PF, elasticita substitúcie sa rovná 0, dopyt po pridanej hodnote je riešený na základe nákladovej minimalizácie CES produkčnej funkcie, pričom vstupy pre pridanú hodnotu sú kapitál (K) a práca (L). Dopyt po jednotlivých komoditách medzispotreby je riešený na základe nákladovej minimalizácie Leontieffovej PF. Štruktúru modelu zobrazuje Obrázok 1.

Obrázok 1: Štruktúra CGE modelu⁶



Rovnice nulového zisku definujú rovnosť medzi príjmami a nákladmi jednotlivých agentov ekonomiky a produkcie. Rovnice trhového vyčistenia definujú rovnosť medzi ponukou a dopytom na príslušných trhoch. Na rozdiel od štandardných CGE modelov model pre vyhodnocovanie NSRR obsahuje dodatočné rovnice, predovšetkým na zmenu technologických parametrov PF a pôsobenie nezamestnanosti na cenu práce. Tieto rovnice vyplývajú z cieľa, na ktorý sa model konštruje a zároveň sú prebrané z konštrukcie modelu HERMIN. V prípade potreby, parametre týchto funkcií boli odhadnuté ekonometricky. Uzavretie modelu je hybridné (obsahuje aj prvky klasického uzáveru aj Keynesovského), teda predpokladá sa nezamestnanosť a teda celkové množstvo práce nie je obmedzené, rovnako objem kapitálu nie je fixovaný ako aj investície. Ako numeraire je fixovaná cenová hladina spotreby domácností (spotrebiteľské ceny).

Hodnotenie jednotlivých scenárov NSRR

⁶ Zdroj: CGE model of the EU structural funds' impacts on the Slovak economy, J. Ďuraš, K. Bakošová, S. Šekereš, Autor modelu: Mgr. Ján Ďuraš

V tejto časti sa budeme venovať vyhodnoteniu jednotlivých scenárov NSRR. Hodnotenie dopadov alokácii NSRR znamená porovnať výsledky vývoja ekonomiky pri použití tohto hospodárskeho nástroja (jednotlivé scenáre) a pri jeho nepoužití čo predstavuje základný (baseline) scenár vývoja ekonomiky bez šokov vyvolaných NSRR. Toky prostriedkov, ktoré predstavujú jednotlivé alokácie v rámci NSRR spôsobujú v ekonomike jednak dopytový šok, ktorý má za následok dočasné zvýšenie rastu ekonomiky, ale zároveň pri správnom využití pôsobia aj na produkčné schopnosti ekonomiky. Hlavným cieľom NSRR je práve zvýšiť produkčné schopnosti ekonomiky. Toto zvýšenie je predovšetkým dosiahnuté akumuláciou kapitálu a to jednak fyzickej infraštruktúry ($KGINF_t$) a zároveň aj kumuláciou ľudského kapitálu ($NTRAIN_t$). Tieto pôsobia v ekonomike na rast produkcie, na základe ich pôsobenia v produkčných funkciách do ktorých sa vloží výraz:

$$\eta_1 \log(KGINF_t / KGINF_0) + \eta_2 \log(NTRAIN_t / NTRAIN_0) \quad (2)$$

Elasticity η_1 a η_2 sú odhadnuté, resp. zvolené podľa štúdie⁷ a to v intervale 0 až 0,4, pričom ak zvolíme nulu nepredpokladáme rast produkčných schopností ekonomiky a celkový efekt alokácii NSRR je redukovaný na dopytový. $KGINF_0$ a $NTRAIN_0$ predstavujú hodnoty zásoby fyzickej infraštruktúry v bázičkom scenári, teda v scenári, v ktorom nepredpokladáme pôsobenie NSRR a ekonomika sa vyvíja bez akýchkoľvek zásahov z tohto prerozdelenia. Ďalší efekt, ktorý môže byť vyvolaný NSRR sú produkčné externality výrobných faktorov, pôsobiace predovšetkým na zmenu technologických koeficientov PF. Tieto pôsobia obdobne ako rast produkčných schopností ekonomiky a vedú k vyššej produktivite kapitálu, resp. práce. To vedie k nižšej náročnosti produkcie na prácu resp. kapitál, keďže na jednotku produkcie je treba menej práce či kapitálu, a zároveň rast produktivity práce vytvára priestor pre rast miezd.

Scenáre, ktoré boli hodnotené týmito modelmi sú „Stredný variant MVRR, august 2005“, „Variant MVRR, máj 2006“ a posledný je „Variant MVRR, október 2006“. Prvé dva varianty boli zostavené vládou pod predsedníctvom Mikuláša Dzurindu, posledný z októbra 2006 súčasťou vládou predsedu Róberta Fica. Výška celkových alokácii sa mierne menila, najdôležitejšia však ostáva zmena štruktúry využitia jednotlivých zdrojov. Počas roku 2005 a 2006 Ekonomický ústav SAV pôsobil ako hodnotiteľ predkladaných scenárov, pričom popri hodnotení bolo snahou kolektívu hodnotiteľov vypracovať aj stručné doporučenia a návrhy zmien v rozdelení alokácii. Posledný scenár na vyhodnotenie alokácii nebol kontrahovaný vládou, napriek tomu sme hodnotenie vykonali pre potreby porovnania s predchádzajúcimi. V prípade scenárov z augusta 2005 a mája 2006 je vidno jednoznačný posun alokácii v prospech špecifických priorít Moderné vzdelávanie pre znalostnú spoločnosť, Rozvoj informačných služieb pre občanov a podniky (informatizácia spoločnosti) a Infraštruktúra

⁷ Vykonané empirické analýzy (napr. ESRI, (1997); Denny, Harmon and O'Connell (2000))

vzdelávania (Infraštruktúra vysokých škôl). Napriek tomu, že súčasná vláda deklarovala vo svojom programovom vyhlásení v odseku 5. Vedomostná spoločnosť že: „**Vláda považuje formovanie vedomostnej spoločnosti za svoju prioritu...**“⁸ a ďalej „**Z uvedeného dôvodu bude podporovať vzdelávacie, vedecko-výskumné, informačné a inovačné aktivity škôl, pracovísk SAV, súkromných a firemných vývojových pracovísk, podnikateľských subjektov s väzbou na verejné a súkromné zdroje, vytváranie lepších legislatívnych podmienok, ako aj na zdroje v rámci nového Národného strategického referenčného rámca 2007 – 2013, osobitne z operačných programov znalostná ekonomika, infraštruktúra vzdelávania, vzdelávanie, regióny a životné prostredie**“ práve v menovaných prioritách urobila najväčšie škrtý a tieto prostriedky presunula v prospech iných priorít. Pre bližšie podrobnosti ohľadne alokácii v jednotlivých scenároch pozri Tabuľka 1.

Tabuľka 1 Alokácie zdrojov z prostriedkov EU v rámci NSRR 2007-2013, jednotlivé varianty, mil.EUR

Špecifická priorita	Alokácie špecifických priorít (operačných programov)		
	Stredný variant MVRR, august 2005	Variant MVRR, máj 2006	Variant MVRR, október 2006
Dopravná infraštruktúra	2 825.9	3 437.0	3 206.9
Environmentálna infraštruktúra a ochrana životného prostredia	2 250.9	1 450.0	1 800.0
Lokálna infraštruktúra (regionálna infraštruktúra)	1 728.0	1 475.0	1 445.0
Podpora rastu konkurencieschopnosti podnikov a služieb prostredníctvom inovácií	1 085.0	850.0	772.0
Výskum, technologický rozvoj a inovácie	859.6	900.0	683.0
Rozvoj informačných služieb pre občanov a podniky (informatizácia spoločnosti)	393.8	850.0	993.1
Podpora rastu zamestnanosti a sociálnej inklúzie	435.2	600.0	864.0
Moderné vzdelávanie pre znalostnú spoločnosť	344.6	800.0	600.0
Zdravá pracovná sila (Modernizácia zdravotníctva)	123.1	200.0	250.0
Infraštruktúra vzdelávania (Infraštruktúra vysokých škôl)	x	575.0	200.0
Technická pomoc	x	100.0	97.6
Rezerva	203.8	x	x
Spolu	10 250.0	11 237.0	10 911.6

Výsledky prepočtu dopadov jednotlivých scenároch pomocou týchto dvoch modelov sú v nasledujúcich tabuľkách (Tabuľka 2, Tabuľka 3, Tabuľka 4). Napriek tomu, že jednotlivé modely sú postavené na rozdielnych princípoch, výsledky prepočtu jednotlivých scenárov sú porovnateľné. Väčší rozdiel pri získaných výsledkoch je predovšetkým v raste produktivity práce, kde pre posledné 2 scenáre pomocou modelu HERMIN dostávame výrazne vyššie dodatočné tempá rastu. V prípade zamestnanosti sú získané výsledky vzájomne porovnateľné a odchýlky sú v roku 2010 nepatrné. Každopádne keďže podľa výsledkov získaných modelom HERMIN sú dodatočné tempá rastu produktivity práce pre ďalšie obdobia vyššie ako

⁸ Bližšie pozri „Programové vyhlásenie vlády Slovenskej republiky“, august 2006, strana 34

výsledky získané modelom CGE, narastá pre roky 2013 až 2015 rozdiel v dodatočnom príraste tempa rastu zamestnanosti. Tento nárast korešponduje s ekonomickou teóriou a rovnako aj s predpokladmi konštrukcie modelu. Rozdiely v dodatočnom tempe rastu HDP pre jednotlivé scenáre sú výrazné predovšetkým medzi prvým a poslednými dvoma scenármi. Odklon od podpory vzdelávania, vedy a výskumu, či vzdelávacej infraštruktúry v poslednom scenári má za následok mierny pokles dodatočného rastu HDP oproti variantu MVRR z mája 2006. Podľa multiplikátora NRR, ktorý poukazuje na multiplikačné efekty jednotlivých variantov, má najlepší navrhnutú štruktúru rozdelenia prostriedkov variant z mája 2006, ktorý generuje najvyššie multiplikačné efekty. Scenár z mája 2006 zaostáva za scenárom z októbra 2006 pri hodnotení CGE modelom v období do roku 2010, čo môže súvisieť s výraznejším dopytovým šokom v počiatočnej fáze čerpania. Pri všetkých ostatných porovnávaných hodnotách vykazuje lepšie výsledky. Scenár z októbra 2006 dosahuje nižšie dodatočné tempa rastu HDP, ako aj multiplikátor NRR pre roky 2013 aj 2015.

Tabuľka 2 Stredný variant MVRR, august 2005

	Hodnotenie modelom HERMIN			Hodnotenie CGE modelom		
	2010	2013	2015	2010	2013	2015
Dodatočný rast HDP	1,4	1,4	1,1	1,5	1,6	1,5
Dodatočné tempo rastu zamestnanosti	0,6	0,5	0,4	0,6	0,5	0,5
Dodatočné tempo rastu prod. práce	0,8	0,9	0,7	1,0	1,0	0,8
Multiplikátor NRR ku koncu obdobia	0,629	0,956	1,158	0,663	1,093	1,519

Tabuľka 3 Variant MVRR, máj 2006

	Hodnotenie modelom HERMIN			Hodnotenie CGE modelom		
	2010	2013	2015	2010	2013	2015
Dodatočný rast HDP	2,0	2,3	2,0	1,7	1,9	1,8
Dodatočné tempo rastu zamestnanosti	0,7	0,4	0,3	0,7	0,6	0,5
Dodatočné tempo rastu prod. práce	1,4	1,9	1,7	1,0	1,2	1,1
Multiplikátor NRR ku koncu obdobia	0,958	1,534	1,852	0,814	1,267	1,670

Tabuľka 4 Variant MVRR, október 2006

	Hodnotenie modelom HERMIN			Hodnotenie CGE modelom		
	2010	2013	2015	2010	2013	2015
Dodatočný rast HDP	1,8	2,1	1,8	1,7	1,8	1,7
Dodatočné tempo rastu zamestnanosti	0,6	0,4	0,3	0,6	0,7	0,5
Dodatočné tempo rastu prod. práce	1,2	1,7	1,6	1,0	1,1	0,9
Multiplikátor NRR ku koncu obdobia	0,891	1,399	1,692	0,841	1,199	1,598

Záver

Historický vývoj alokovaných prostriedkov pre jednotlivé verzie NSRR bol v procese vyhodnocovania jeho dopadov rozdielny. Z predložených výsledkov je však zrejmé, že rozhodujúci vplyv pre posúdenie efektívnosti ich využitia má štruktúra použitia týchto

prostriedkov ako ich celkový objem.. Výrazné zmeny v dodatočnom raste produktivity práce či rastu HDP sa dajú dosiahnuť práve vďaka optimalizovaniu štruktúry týchto alokácií. Na základe získaných výsledkov konštatujeme, že svojou štruktúrou sa práve variant z mája 2006 naviac priblížil k tomuto optimálnemu rozdeleniu. Tento variant predstavuje výrazný pozitívny posun oproti prvému z augusta 2005, teda zmena rozdelenia bola tvorcami posunutá správnym smerom. Posledný variant z októbra 2006 oproti tomu dosahuje podľa hodnotenia modelom HERMIN aj CGE modelom horšie výsledky. Príčinou toho je práve zmena štruktúry použitia týchto prostriedkov. Tento variant teda predstavuje odklonenie od efektívneho vektora. Faktom naopak zostáva, že scenár z októbra 2006 nebol konzultovaný s partnerskými organizáciami ani hodnotiteľom (EÚ SAV) a ani nebolo požadované hodnotenie. Na druhú stranu, štruktúrne modely ako je model HERMIN či CGE model sa ukazujú byť vhodnými nástrojmi na hodnotenie NDP. Na základe získaných výsledkov, pre rôzne scenáre a s rôznou štruktúrou, môžeme jednoznačne konštatovať, že tieto modely výborne plnia svoju úlohu ako jeden z možných nástrojov hodnotenia.

Literatúra

BRADLEY J., MODESTO L., SOSVILLA-RIVERO S. 1995, HERMIN, A macroeconometric modelling framework for the EU periphery, Economic Modeling, č. 3, 1995, s. 221-247

BRADLEY J., MODESTO L., SOSVILLA-RIVERO S. 1995, Similarity and diversity in the EU periphery, A HERMIN – based investigation, Economic Modeling, č. 3, 1995, s. 313-322

VAŇO B. 2004, Prognóza vývoja obyvateľov v okresoch SR do roku 2025, Bratislava, Infostat, 2004

BLANCHARD O. J., QUAH D. 1989, The Dynamic Effects of Aggregate Demand and Supply Disturbances, In: The American Economic Review, 1989, č. 4, s. 655-673,

LOFGREN H., HARRIS LEE R., ROBINSON S. 2002, A standard computable general equilibrium (CGE) model in GAMS, Washington, D.C.: Intl Food Policy Research Inst, 2002. s.79, ISBN 0-89629-720-9

Adresa:

Ekonomický Ústav SAV
Šancová 56
811 05 Bratislava
martin.mlynek@savba.sk

Regionálna prognóza SR so zameraním na trh práce a HDP do roku 2017

Vladimír Kvetan, Marek Radvanský¹

Abstract

There is a wide discussion regarding the regional disparities during the past years in Slovakia.. To understand, estimate and forecast future trends and possible development of Slovak regions it is necessary to use regional model. Described model was build at Institute for Economic Research at Slovak Academy of Sciences in Bratislava, using econometric and expert approaches. The emphasis is on development of labor market and production of domestic product. The presented model is concentrated on forecasting the labor market development and GDP until 2017 up to level LAU1.

Key words: regional model, econometrics, labor market, regional disparities

Úvod

Regionálny model Slovenska bol vytvorený na Ekonomickom ústave SAV v nadväznosti na makroekonomický model Slovenska typu ECM_BIER 07Q1. Vzhľadom na dostupnosť regionálnych dát bol zameraný na očakávaný vývoj hlavne v oblasti trhu práce. Model je založený na ročných dátach v územnom členení podľa NUTS2, NUTS3 a LAU1². Ako vstupné dáta boli použité regionálne údaje za trh práce, produkciu podľa sektorov a HDP v členení podľa NUTS3. Model bol skonštruovaný tak, aby umožnil prognózovať indikátory trhu práce ako na základe štatistickej evidencie, tak i VZPS. Základ prognózy vývoja obyvateľstva bola prognóza Výskumného demografického centra (Vaňo (2004)). V prípade detailnejšej prognózy na úrovni okresov boli použité dostupné údaje do roku 2004.

Pri tvorbe regionálnej prognózy tvorila základ makroekonomická prognóza Slovenska, ktorá vznikla na základe použitia modelu ECM_BIER_07Q1³. Expertné metódy vychádzali z troch základných princípov – analýza východiskového stavu; analýza potenciálu; analýza budúceho rozvoja. Hlavnou myšlienkou regionálnej prognózy bol predpoklad uplatňovania takej regionálnej politiky, ktorá by viedla k postupnému zastavovaniu narastania regionálnych disparít. Na druhej strane by však táto regionálna politika nemala brániť rastu celej národnej ekonomiky SR a jej postupnej konvergenie k priemeru EU. Vznik regionálnych disparít nie

¹ Ekonomický ústav SAV, marek.radvansky@savba.sk, vladimir.kvetan@savba.sk

² úroveň „starých krajov“ cez VUC až po úroveň okresov

³ pozri Pohľady na ekonomiku Slovenska, 2007

je krátkodobou záležitosťou a preto ani ich odstránenie by nemalo byť krátkodobým cieľom. Aj pri porovnaní ostatných európskych krajín sú zjavné väčšie, alebo menšie regionálne nerovnosti. Ich hlavnou príčinou sú popri ekonomických danostiach aj geografické, demografické a sídelné charakteristiky. Je teda pochopiteľné, že predkladaná prognóza neuvažuje s rýchlym riešením problému regionálnych disparít.. Teoretickým východiskom je rastovo – egalizačný prístup, ktorý vychádza predovšetkým z podpory rozvojových centier pri súčasnom dobudovávaní infraštruktúry v zaostalých častiach regiónov. Do modelu boli zapracované predpoklady o konkrétnych potvrdených investíciách v jednotlivých regiónoch ku koncu roka 2006, ako aj predpoklady zohľadňujúce aktuálny návrh Národného strategického referenčného rámca⁴.

Na odhady jednotlivých závislostí v rámci krajov a okresov bol použitý ekonometrický softvér E-Views®. Následne bola pomocou riešenia multikriteriálnej úlohy nástrojom Solver v MS Excel® docielená medzi jednotlivými hodnotami horizontálna konzistentnosť (medzi regiónmi), ako aj vertikálna konzistentnosť (medzi jednotlivými makroindikátormi). Ako už bolo spomenuté prognóza je vyčíslená až na úroveň okresov, čo je možné označiť ako unikátne vzhľadom na slovenské prognostické pracoviská.

Východiskový stav v krajoch NUTS II a NUTS III

Značne nevyrovnaný vývoj ekonomickej výkonnosti (a nadväzujúcich charakteristík) v regiónoch NUTS II je spojený s ako historickými danosťami tak aj s doterajšou ekonomickou politikou:

- Charakter industrializácie týchto regiónov viedol v minulosti ku vzniku monoštruktúrnych mikroregiónov, v ktorých rozhodujúcim zamestnávateľom (a tvorcom pridanej hodnoty) bol obyčajne malý počet veľkých závodov. Tým sa regióny stali veľmi zraniteľné. Po liberalizácii trhov nosné podniky rýchlo stratil svoju konkurencieschopnosť. Náhrady týchto rozhodujúcich zamestnávateľov však vznikajú iba postupne.
- V centrálne plánovanej ekonomike pri existencii silnej redistribúcie nevznikali veľké príjmové disproporcie napriek rozdielnym výsledkom v produktivite či tvorbe pridanej hodnoty (ktorá však nebola ani rozhodujúcim ukazovateľom pri prevládaní fyzických ukazovateľov materiálnej výroby). Po zmene vlastníckych vzťahov a zmene redistribučného mechanizmu sa však diferencie rýchlo vyostрили bez toho, aby bola ekonomická a sociálna politika pripravená ich zmierňovať.

⁴ Konečná verzia NSRR stále neexistuje

- Direktívne riadená lokalizácia dôležitých podnikov viedla k neprirodzenému umiestneniu energeticky, environmentálne náročných výrobných s nízkou tvorbou pridanej hodnoty do regiónov, ktoré bez pomoci centrálného plánovania neboli schopné zabezpečiť potreby pre ďalší rozvoj alebo aspoň udržanie existujúceho stavu. Nutnosť reštrukturalizácie takýchto výrobných znamenala vážny dopad na zamestnanosť a tvorbu zdrojov v týchto regiónoch..
- Doterajšia regionálna politika v SR bola založená skôr na proklamatívnych vyhlásení, než na reálne uskutočňovaných krokoch vedúcich k zmierňovaniu regionálnych disparít.
- Z pozitívnych efektov cezhraničnej kooperácie s vyspelejšími regiónmi mohlo čerpať iba západné Slovensko. Stredné a východné Slovensko hraničí s málo vyspelými regiónmi Maďarska, Poľska a Ukrajiny s malou šancou na tvorbu silných klastrov čo takisto prispelo k prehĺbovaniu rozdielov..

Charakteristiky podľa jednotlivých NUTS II:

Bratislavský kraj: Je jediným krajom SR, ktorého HDP na obyvateľa dosahuje úroveň priemeru bývalej EU 15. Zároveň je jediným regiónom s nedostatkom pracovných síl, preto je cieľom migrácie za prácou. Je typickým mestským krajom, ktorý je problematické porovnať s ostatnými krajinami, ktoré majú rozdrobenú sídelnú štruktúru s prevahou rurálneho osídlenia. Bratislavský kraj má preto celkom pochopiteľne trojnásobne vyššiu úroveň HDP na obyvateľa v porovnaní s ostatnými krajinami NUTS II, viac ako päťnásobne vyššiu mieru produktivity práce v priemysle, o 80% vyššiu úroveň priemernej nominálnej mzdy v porovnaní s priemerom SR. Kraj navyše ťaží zo svojej excentrickej polohy, keď je „vysunutý“ do priestoru v blízkosti iných výrazne prosperujúcich miest, ako je Viedeň alebo Győr. Vzniká tak geografická blízkosť troch silno rastúcich regiónov, ktoré sa vo svojom raste vzájomne podporujú. Excentrická poloha kraja je však nevýhodou pre ostatné kraje Slovenska. Mestský charakter Bratislavského kraja dáva predpoklady pre jeho silný rast prakticky pri rôznej konštelácii hospodárskej politiky, dáva to kraju vysokú mieru odolnosti voči externým ekonomickým šokom. V Bratislavskom kraji je najvyššia koncentrácia výrobných s vysokou pridanou hodnotou, jedine tento kraj začína plniť funkciu technologického lídra.

Západoslovenský kraj: Má možnosť ťažiť z relatívnej blízkosti k vyspelým susediacim regiónom (Bratislavský kraj, Česká republika, severozápad Maďarska). V tomto regióne sa prejavuje diferenciácia ekonomickej výkonnosti medzi rýchlo rastúcim pásom rovnobežným s hranicami s ČR (koridor okolo diaľnice D1, zahŕňajúci Trnavu, Ilavu, Trenčín, Púchov) a zaostávajúcim juhovýchodným regiónom. Táto diferenciácia má tendenciu narastať, je umocnená prílevom PZI do západnej a severozápadnej časti regiónu a naopak

periférnym charakterom juhovýchodnej časti. Okresy juhovýchodnej časti regiónu sú poznačené typickou kombináciou agrárneho vidieka a mesta s malým počtom nosných priemyselných činností, s nízkou mierou diverzifikácie. Kríza priemyselných závodov v mestách v procese vlastníckych zmien a reštrukturalizácie viedla k úpadku rozhodujúcej priemyselnej aktivity (prípady Novofruct Nové Zámky, Elektrosvit Nové Zámky, Slovenské lodenice Komárno, Plastika Nitra a pod.). Niektoré z nosných priemyselných činností sa znovu oživilo, ale nedokázali nahradiť rozsah pôvodnej zamestnanosti. Výraznou charakteristikou južnej, zaostávajúcej časti regiónu je aj jej demografická depresia.

Stredoslovenský kraj: Je tiež diferencovaný severo - južne: výkonnejší sever (okrem krajného severu ako Orava a Kysuce) a stred má ekonomické parametre porovnateľné s priemerom SR, južný pás je typickým periférnym regiónom. V južnom páse, podobne ako v prípade juhovýchodu Západoslovenského kraja, platí kombinácia agrárneho charakteru s málo diverzifikovanými nosnými priemyselnými aktivitami (a úpadkom týchto bývalých nosných priemyselných aktivít v priebehu transformácie ekonomiky). Značný rozsah sociálnej exklúzie a osôb s absenciou pracovných návykov na južnej periférii kraja vytvára dopyt po opatreniach zameraných na kohéziu a hlavne sociálnu inklúziu. Kraj je značne postihnutý (podobne ako východoslovenský) rozdrobenosťou sídelnej štruktúry s vysokým podielom malých obcí. Funkciu generátorov rastu spĺňajú predovšetkým mesto Žilina (ktorého okolie je pokračovaním rozvojového pásu Bratislava – Trnava- Trenčín) a mikroregión miest Zvolen-Banská Bystrica. Sila týchto pólov rastu však vzhľadom na rozmery kraja a jeho rozdrobenú sídelnú štruktúru nie je dostačujúca.

Východoslovenský kraj: Ide o kraj, kde nie je nápadná severojužná diferenciácia ako v iných krajoch, ale skôr diferenciácia medzi relatívne výkonným pásom Prešov – Košice a zaostávajúcim zvyškom regiónu, ktorý má najmenej priaznivé socio-ekonomické parametre v SR. Severná časť regiónu (Prešovský kraj) má najmenší podiel vysokoškolsky vzdelaného obyvateľstva a najvyšší podiel obyvateľov vyučených v robotníckych profesiách v SR. Tento región má nízky index starnutia a najvyšší populačný prírastok v SR. Ťahúň južnej časti, mesto Košice, opticky výrazne vylepšuje štatistické údaje regiónu, toto mesto však zatiaľ v žiadnom prípade nemá a nemôže mať takú ekonomickú silu, aby zabezpečilo dostatočnú mieru ekonomickej výkonnosti pre celý región. Touto silou zatiaľ nedisponuje ani jeden z úspešných slovenských priemyselných parkov, napr. Kechnec, ktorý je jedným z mála úspešne etablovaných ekonomických aktivít v kraji. Celý východoslovenský región je poznačený mimoriadne silným zastúpením marginalizovaných skupín (predovšetkým rómskeho etnika, ale nielen toho). Chronická silná nerovnováha na trhu práce, silné

zastúpenie marginalizovaných skupín si vyžadujú značné aktivity zamerané na sociálnu inklúziu.

Hoci má región dobrý potenciál pre rozvoj cestovného ruchu, v súčasnom konkurenčnom prostredí v rámci stredoeurópskeho regiónu nie je cestovný ruch odvetvím, ktorý by dokázal garantovať dostatok pracovných miest či dostatočnú tvorbu zdrojov. Je pozoruhodné, že ani v turisticky najatraktívnejších okresoch (tých, ktoré majú na svojom území Tatry – t.j. Poprad a Kežmarok), nie je podiel zamestnanosti v cestovnom ruchu dramaticky vyšší ako v priemere za SR. Cestovný ruch teda zatiaľ nedokáže plniť rozhodujúce rozvojové funkcie regiónu. Jednou z možných prekážok v budovaní cestovného ruchu sa môže javiť nedostatočná dopravná infraštruktúra a úroveň ostatných služieb, ktorá výrazne znižuje atraktivnosť tohto regiónu.

Tabuľka 1 Priemerné hodnoty ukazovateľov (v rokoch 2000 - 2005)

	Tempo rastu HDP s.c.	Podiel HDP v NUTS II na SR	Tempo rastu počtu zamestnaných	Miera nezamest. z disp. počtu nezamest.	Priemer podielu produktivity práce
Bratislavský kraj	5,5%	26,1%	0,9%	4,5%	142,4%
Západné Slovensko	3,6%	31,6%	1,2%	13,8%	94,6%
Stredné Slovensko	4,5%	20,6%	0,3%	17,5%	89,1%
Východné Slovensko	4,3%	21,7%	0,3%	21,2%	86,4%
SR spolu	4,4%	100%	0,7%	15,6%	100%

Zdroj: ŠÚ SR

Spomenuté charakteristiky sa zákonite odrazili aj na makroekonomických ukazovateľoch za jednotlivé NUTS II. Tabuľka 1. zachytáva priemerné ukazovatele za roky 2000 – 2005. V rokoch 2000 – 2005 bol najrýchlejšie rastúci Bratislavský kraj (5,5%) a najpomalším krajom západné Slovensko. Vyššie priemerné tempá rastu vo Východoslovenskom a Stredoslovenskom kraji boli zaznamenané najmä z dôvodu nižších porovnávacích základov. Je to badateľné z jednotlivých príspevkov krajov na HDP: Najvyšší priemerný podiel má Západoslovenský kraj a najnižší Stredoslovenský kraj. Západoslovenský kraj tvoril aj najvyšší prírastok v evidovanom počte zamestnaných. Bratislavský kraj zaznamenal v tomto období mierny rast priemerného počtu zamestnaných o 0,9%. Aj napriek tomuto miernemu rastu dosahuje Bratislavský kraj najnižšiu mieru nezamestnanosti. Naopak stabilne je najvyššia miera nezamestnanosti na východe Slovenska. Rovnaké charakteristiky dosahuje aj podiel priemernej produktivity práce v jednotlivých krajoch k priemeru za SR: Z tabuľky je zrejmé, že Bratislavský kraj dosahuje vysoko nadpriemernú hodnotu tohto ukazovateľa (142,4%).

Predpoklady o budúcom vývoji v regiónoch NUTS II

Vývoj v jednotlivých NUTS II regiónoch bude založený na ich súčasnom východiskovom stave v súčinnosti s predpokladanou politikou rozvoja. Predpokladáme priaznivý efekt prílevu priamych zahraničných investícií (PZI), úspešnú aplikáciu regionálnej politiky a takisto podporu z NSRR kofinancovaného z fondov EÚ spolu s verejnými zdrojmi. Význam tejto podpory bude nielen v samotnom rozvoji výroby, výstavby a tvorby pracovných miest. Vysoko pozitívne budú pôsobiť následné synergické a multiplikačné efekty. Výrazným prvkom bude nevyváženosť budúceho vývoja na úrovni NUTS II. Je vysoko pravdepodobné, že kraje (úroveň NUTS III) s relatívne pozitívnym vývojom budú brzdené krajinami s horším vývojom. Výsledkom toho budú aj v budúcnosti niektoré regióny podhodnotené, iné naopak nadhodnotené. Príkladmi sú kraje Nitriansky a Banskobystrický, ktoré vplyvom vykazovania v rámci NUTS II sa javia ako bezproblémové, na druhej strane je ich vývoj hlboko podpriemerný.

Bratislavský kraj – svojou polohou, východiskovým stavom a statusom hlavného mesta je tento kraj predurčený aj naďalej k nadpriemerne pozitívnemu vývoju. V budúcnosti predpokladáme pretrvávajúce výrazné prílevy PZI spolu s kumuláciou investícií z národných zdrojov. Mestský charakter tohto regiónu ho predurčuje aj naďalej ako centrum hospodárskeho, finančného a vedecko-technologického progresu. Spolu s postavením hlavného mesta a teda centra verejnej správy to bude pozitívne vplývať na tvorbu pracovných miest s vysokou produktivitou práce. Nebratislavské okresy Bratislavského kraja budú profitovať z ich polohy v bezprostrednej blízkosti centra. Podstatným faktorom rozvoja týchto okresov bude aj naďalej pretrvávajúci trend sťahovania sa z mesta na vidiek a vznik nadväzných služieb – okresy Pezinok, Malacky a Senec.

Západoslovenský kraj – významným faktorom rozvoja tohto kraja bude jednak jeho poloha relatívne blízko centra (Bratislavy) spolu s hranicou z ČR a tiež vysoký potenciál rozvoja kde sa spája kvalifikovaná pracovná sila s tradíciou strojárenského, spracovateľského a iného priemyslu. Z blízkosti centra bude významne ťažiť Trnavský samosprávny kraj ktorý stále tvorí ponuku práce pre Bratislavský kraj. Južné okresy Trnavského kraja budú tiež profitovať z možnosti práce v susedných severných rozvinutých regiónoch Maďarska. Trnavský kraj bude v budúcnosti taktiež profitovať aj z výstavby automobilového závodu, kde sa dá očakávať rozvoj nadväzujúcich výrobných a rast dopytu po službách v okolí tejto investície. Trenčiansky kraj bude ťažiť so súčasného významného postavenia v rámci SR. Výhodná je i poloha v blízkosti hraníc s ČR a medzi dvoma krajinami s rozvinutou priemyselnou štruktúrou –

Trnavským a Žilinským. Je vysoký predpoklad, že nadväzná výroba a služby pre oba automobilové komplexy, ktoré je možné využívať spoločne budú alokované práve do tohto kraja. Vysoko problémovým sa do budúcnosti javí vývoj v Nitrianskom kraji. Jeho súčasná rozvinutosť nezakladá dostatočné predpoklady na výraznejšie zlepšenie. Najmä v južných okresoch s prevažujúcou poľnohospodárskou výrobou nie je možné očakávať výraznejšie zlepšenie. Problematická bude aj podpora tohto regiónu z fondov EÚ, nakoľko kritérium finančného prerozdelenia je založené na členení NUTS II. Keďže sa tento kraj nachádza v skupine s výrazne silnými regiónmi jeho celková pozícia je výrazne nadhodnotená a preto sa dostáva na relatívny okraj záujmu podpory.

Stredoslovenský kraj – Budúci vývoj tohto kraja bude výrazne nerovnomerný. Žilinský kraj bude významne profitovať v najbližších rokoch z prílevu PZI a plánovanej výstavbe automobilových závodov. Takisto predpokladáme pokračujúcu prosperitu v súčasných rozvinutých centrách – okresy Ružomberok, Martin, Liptovský Mikuláš. Problematický vývoj očakávame v Oravskom regióne. V Banskobystrickom kraji v budúcnosti očakávame značne nevyvážený rozvoj. Výrazným sa javí severo – južný problém, kde v severných okresoch⁵ je produkcia orientovaná prevažne na priemyselnú výrobu (drevárstvo, strojárstvo, spracovanie kovov). Naopak južné okresy⁶ sú prevažne poľnohospodárske. Do budúcnosti nepredpokladáme výraznú zmenu tohto pomeru. Tradícia a kvalifikovanosť pracovnej sily bude výrazne pôsobiť v prospech rozvoja severných okresov. Naopak južné okresy budú aj naďalej relatívne zaostávať. Z hľadiska vývoja na trhu práce je možná i pracovná migrácia obyvateľstva do severných okresov Banskobystrického kraja tohto kraja.

Východoslovenský kraj – Aj napriek výraznej podpore zo strany eurofondov ktorú v tomto kraji v rokoch 2006 – 2013 predpokladáme, nebude vývoj v tomto kraji v strednodobom horizonte tejto prognózy výrazne pozitívny. Košický kraj bude významne čerpať z rozvoja v centre – mesto Košice. To bude pôsobiť ako rozvojové centrum a tiež ako hlavný zamestnávateľ v kraji. Rozvoj Prešovského kraja bude výrazne závislý od budovania technickej infraštruktúry. Výrazné zaostávanie je spôsobené najmä súčasnou slabou rozvinutosťou celého Východoslovenského kraja. Slabá rozvinutosť technickej infraštruktúry znemožňuje využívanie na rozvoj priemyslu, služieb a cestovného ruchu. Významným zdrojom nerovnováh bude aj v budúcnosti demografická skladba a populačný vývoj obyvateľstva.

⁵ Banská Bystrica, Zvolen, Žarnovica, Banská Štiavnica, Detva

⁶ Lučenec, Veľký Krtíš, Rimavská Sobota, Poltár,

Prognóza socio-ekonomického vývoja v regiónoch NUTS II a NUTS III do roku 2017

Očakávaný socio-ekonomický vývoj v hlavných regiónoch SR na úrovni NUTS II bude korešpondovať z makroekonomickou prognózou SR pre roky 2007-2017. Očakáva sa vyššia dynamika rastu národného produktu v symbióze s očakávaným poklesom nezamestnanosti. Regionálna politika bude pokračovať v snahe znižovať regionálne disparity na všetkých úrovniach, čím sa dosiahne výraznejšie spomalenie divergenčného procesu.. Napriek tomu však očakávame aj naďalej mierny nárast rozdielov medzi jednotlivými regiónmi, a ku konvergencii bude dochádzať len v rámci vybraných menších regiónov - okresov. Napríklad mzda v Bratislavskom kraji sa bude i naďalej vzdalovať od zbytku Slovenska, ale v jednotlivých okresoch Bratislavy sa budú mzdové rozdiely vyrovnávať⁷.

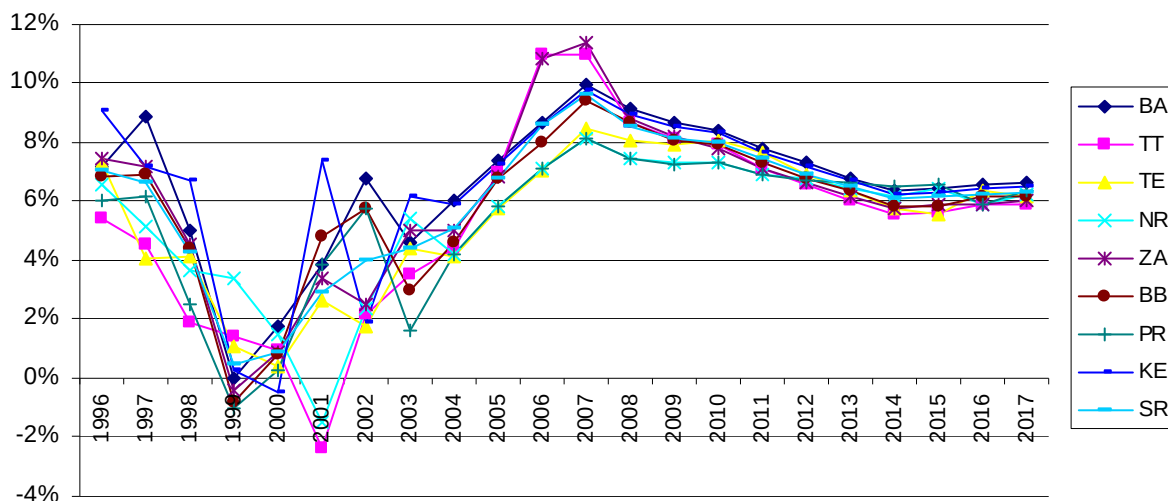
HDP

Pre Slovensko bude v rokoch 2007-2010 očakávaný priemerný rast HDP v stálych cenách na úrovni 8,6 %, pričom najvyššie rasty očakávame v prvých rokoch prognózy v kontexte s prejavom efektov získaných PZI v predchádzajúcich obdobiach. Vysokú priemernú dynamiku rastu si v tomto období udržia hlavne silnejšie kraje (Bratislavský, Trnavský, Žilinský a Košický) na úrovniach okolo 0,5% bodu nad celoslovenským priemerom. Bratislavský kraj s priemerným rastom na úrovni 9 % zároveň považujeme za jediný región schopný výraznejšie generovať domáce investície. Z dôvodu predpokladaného výrazného prílevu PZI bude v porovnaní ostatných krajov rásť najrýchlejšie stredoslovenský kraj a to najmä v úvodných rokoch prognózy majoritne príspevkom Žilinského kraja. Výraznejšie bude zaostávať Nitriansky a Prešovský kraj s rastami okolo 1% bod pod Slovenským priemerom. V strednodobom a dlhodobom horizonte na roky 2011-2017 očakávame vyrovnanější vývoj v regiónoch na priemernej úrovni 6,5 %. Dynamika rastu HDP v západoslovenskom regióne bude mierne brzdená vplyvom miernejšieho rastu produktivity práce vzhľadom na relatívne dobrý východiskový stav a na momentálne vysoko podpriemerné regióny stredného a východného Slovenska. predpokladáme udržiavanie pomerne vysokého rastu HDP vo východoslovenskom kraji vplyvom synergických efektov budovania infraštruktúry a čerpania eurofondov v slabšie rozvinutých regiónoch (najmä Prešovský kraj s nízkym súčasným základom). V Trenčianskom kraji môžeme v strednodobom horizonte očakávať pozitívne efekty vplyvom výhodnej polohy medzi centrami automobilového priemyslu v osi Žilina – Trnava. Ani na úrovni NUTS III nepredpokladáme kraj s výrazne

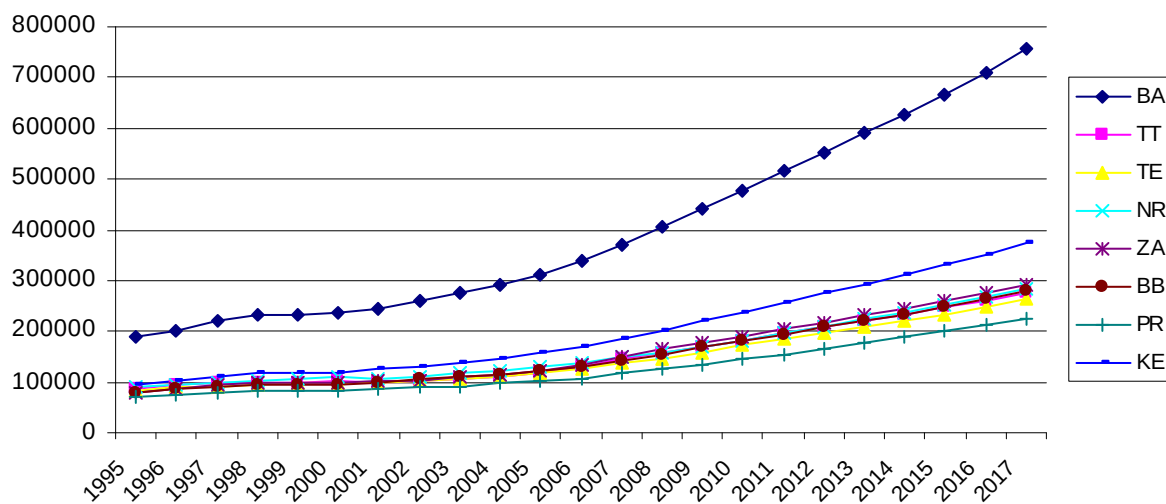
⁷ okrem okresu Bratislava I

podpriemernými rastami produkcie v rámci sledovaného obdobia prognózy. Pri pohľade na vývoj štruktúry produkcie ekonomiky predpokladáme len mierne zmeny v prospech Bratislavského, Žilinského a Trnavského kraja spolu s miernym oslabením ostatných krajov. Disparity sa ani v rámci samotných krajov výrazne neznížia, výnimku očakávame v prípade nezamestnanosti, kde zníženie celkovej nezamestnanosti na Slovensku povedie k znižovaniu rozdielov aj v jednotlivých regiónoch.

Graf 1 – Tempá rastu HDP za NUTS III, s.c.



Graf 2 – Prognóza vývoja HDP za NUTS III, s.c., mil. Sk



Prameň – modelové výpočty tímu autorov

V produktivite práce na zamestnanca v hospodárstve SR je vidno výraznejšie rozdiely v rámci regiónov. Priemerná produktivita práce na Slovensku v bežných cenách vzrastie zo súčasnej hodnoty 780 tisíc Sk v roku 2006 na 1050 tisíc Sk v roku 2010 a 1633 tisíc v roku 2017. Výraznejší predstih Bratislavského kraja pred zbytkom Slovenska je spôsobený

sústredením odvetví z vysokou pridanou hodnotou (bankovníctvo, financie) a zároveň sú tu sústredené sídla väčšiny inštitúcií a nadnárodných spoločností. Dosahuje viac ako 1,5 násobok priemernej produktivity práce na Slovensku a očakávame mierny rast tohto podielu i v ďalších obdobiach⁸. V produktivite práce nasleduje západoslovenský kraj ťahaný hlavne vyššou produktivitou práce v Trnavskom kraji. Ďalej nasleduje stredoslovenský kraj a najslabšie je na tom východné Slovensko. Tu je markantný rozdiel medzi Prešovským a Košickým krajom. Košický kraj pritom dosahuje priemerné hodnoty produktivity práce, ale región je výrazne brzdený Prešovským krajom, ktorý dosahuje len 68 % priemernej produktivity práce bez výraznejších očakávaní k jej vyšším rastom v horizonte prognózy.

Trh práce

Vývoj na trhu práce bude poznačený jednak vývojom populácie a jednak ekonomickou výkonnosťou v regióne. Z hľadiska vývoja populácie bude najrýchlejšie rásť východoslovenský kraj (Vaňo (2004)). Jeho populácia vzrastie z 1,57 mil. v roku 2006 na 1,61 mil. v roku 2013, čo je o takmer 48 tis. obyvateľov viac. Priemerný prírastok v tomto kraji bude na úrovni 5219 obyvateľov ročne. Počet obyvateľov Stredoslovenského kraja bude rásť najmä z dôvodov vyššieho prírastku v oravských okresoch⁹ priemerným ročným prírastkom 942 obyvateľov. V západoslovenskom kraji bude počet obyvateľov klesať v horizonte prognózy o takmer 1500 obyvateľov ročne. Príčinami sú zmena rodinného správania sa, migrácia bližšie k centru (Bratislavský kraj) a do českej republiky. Výrazným faktorom sú negatívne prírastky, ktoré sú charakteristické pre južné okresy Slovenska. Počet obyvateľov Bratislavského kraja bude relatívne stagnovať s prírastkom 144 obyvateľov ročne, pričom absolútny prírastok má tendenciu v čase klesať.

Z celkovej populácie je možné na základe mier participácie odvodiť vývoj ekonomicky aktívneho obyvateľstva. Najvýraznejší pokles ekonomicky aktívneho obyvateľstva bude v Bratislavskom kraji. Medzi rokmi 2006 – 2017 poklesne počet ekonomicky aktívnych o približne 18500 obyvateľov. Rast ekonomicky aktívneho obyvateľstva bude zaznamenaný v regiónoch východného a stredného Slovenska (do roku 2017 to bude 19900, resp. 43600 osôb).

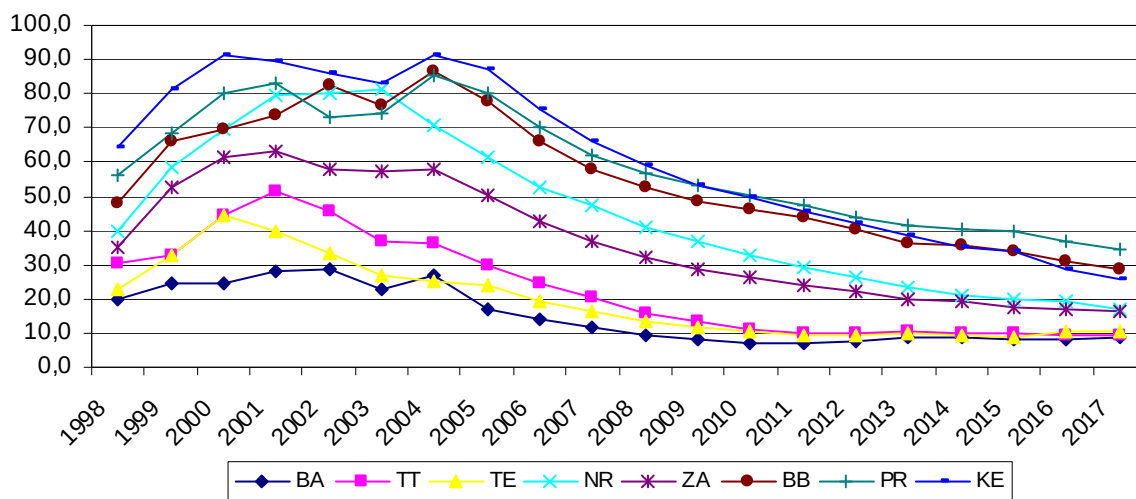
Počet zamestnaných má z regionálneho hľadiska dve roviny. Na jednej strane sú regionálne počty o zamestnaných osobách počítaných na základe výberového zisťovania pracovných síl (VZPS) a jednak na základe evidencie Národného úradu práce (NÚP). Tieto zistenia dávajú prehľad koľko obyvateľov daného regiónu je zamestnaných pričom

⁸ Napriek tomu že produktivita práce na zamestnanca v hospodárstve SR zohľadňuje pravidelnú migráciu za prácou. Napr. do Bratislavy denne dochádza alebo pracuje viac ako 100 tisíc ľudí, ktorí tu nemajú trvalý pobyt.

⁹ Dolný Kubín, Námestovo, Tvrdošín

nezohľadňuje miesto ich zamestnania. Druhá rovina predstavuje evidencia počet zamestnaných v národnom hospodárstve (PZNH), ktorá eviduje počet pracujúcich v regióne, pričom nezohľadňuje ich miesto bydliska.

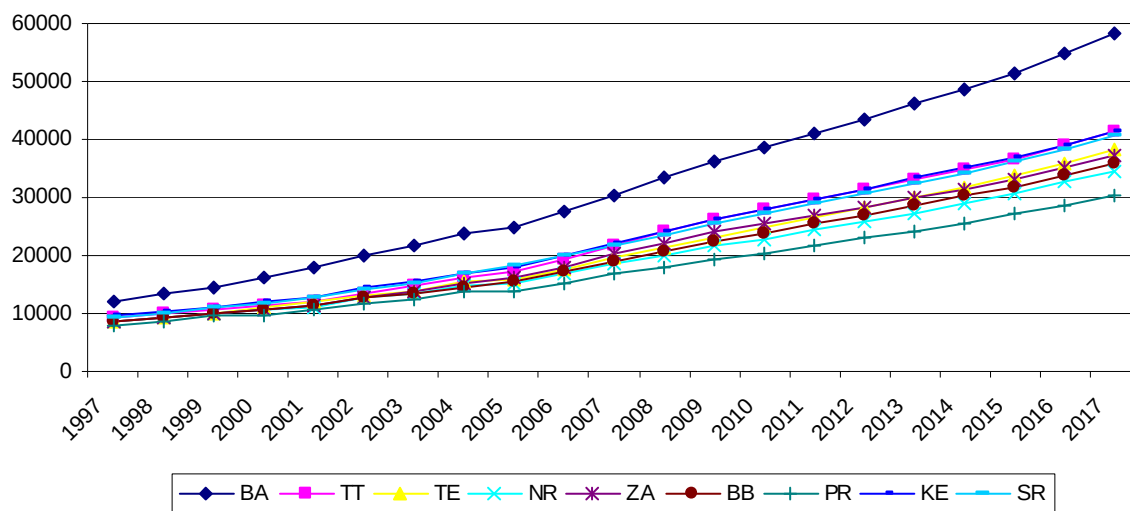
Graf 3 – Prognóza vývoja počtu nezamestnaných v krajoch podľa VZPS



Z hľadiska sledovania zamestnanosti podľa VZPS a podľa evidencie NÚP predpokladáme pomalé približovanie sa týchto dvoch metodík. Na celonárodnej úrovni však toto približovanie bude mať len marginálny efekt.

Vývoj dynamiky počtu zamestnaných bude kopírovať zmeny vo výkonnosti ekonomiky v danom regióne. Vo všetkých krajoch (s výnimkou Bratislavského) bude zaznamenávaný stabilný rast dopytu po práci. Z titulu nášho predpokladu o príleve PZI a takisto mobilizácie domácich zdrojov bude na západnom Slovensku stabilná dynamika rastu zamestnanosti, zamestnanosť na strednom Slovensku bude rásť najrýchlejšie v úvodných rokoch prognózy (2005 – 2008). Vo Východoslovenskom kraji sa bude dynamika rastu zamestnanosti zvyšovať ku koncu prognózy. Zaujímavý bude vývoj v Bratislavskom kraji, kde bude dopyt po práci postupne klesať. Bude to spôsobené postupným zvyšovaním produktivity práce. Nakoľko sa však jedná o relatívne malé úbytky vzhľadom na súčasný stav pracujúcich a takisto sa predpokladá aj postupný pokles počtu ekonomicky aktívneho obyvateľstva nebude mať toto znižovanie výrazný vplyv na trh práce.

Graf 4 – Prognóza vývoja hrubej priemernej mesačnej nominálnej mzdy v Sk



Rozdielny vývoj miery dopytu po práci a ekonomicky aktívneho obyvateľstva v jednotlivých regiónoch bude mať samozrejme aj rozdielny vplyv na mieru nezamestnanosti. Ako bolo možné pri formovaní prognózy očakávať najnižšia miera nezamestnanosti bude dosahovaná v Bratislavskom kraji. Smerom od západu na východ bude miera nezamestnanosti postupne rásť, a najvyššia miera nezamestnanosti bude dosahovaná vo Východoslovenskom kraji. Značný nesúlاد bude však badateľný pri miere nezamestnanosti Vo vnútri NUTS II – na úrovni samosprávnych krajov, prípadne okresov. Napriek tomu, že Stredoslovenský kraj bude ako celok vykazovať prijateľnú mieru nezamestnanosti, práve v okresoch Banskobystrického kraja¹⁰ budú aj naďalej pretrvávajúť najvyššie dosahované miery nezamestnanosti. Z titulu nízkych mier nezamestnanosti v okresoch Žilinského kraja¹¹ nebude tento vývoj na úrovni NUTS II badateľný. Naproti tomu na východe Slovenska bude dosahovaná nižšia miera nezamestnanosti oproti extrémom na strednom Slovensku, avšak bude dosahovaná na vo väčšom počte okresov, čo bude pôsobiť na celkový nepriaznivý obraz o nezamestnanosti v tomto regióne.

Záver

Regionálny model Slovenska je dôležitý nástroj na analýzu postupných zmien v regionálnom vývoji a pri aktívnej snahe znižovať regionálne disparity. Napriek tomu bude mať najvyšší vplyv na vývoj v jednotlivých regiónoch súkromný sektor, a i napriek aktívnej snahe sa budú niektoré regionálne rozdiely i naďalej prehĺbovať. V rámci krajov a okresov bude najväčší rast sústredený v regionálnych centrách a väčších mestách. Príspevok priniesol

¹⁰ Najmä okresy Poltár, Rimavská Sobota, Revúca. Veľký Krtíš

¹¹ Najmä okresy Žilina, Námestovo, Čadca

stručný prehľad o očakávanom vývoji v regiónoch s krátkodobého a strednodobého hľadiska s výhľadom do roku 2017. I keď je táto prognóza tvorená až na úroveň okresov, prezentované boli iba hlavné trendy očakávaného vývoja.

Tabuľka 2 – Vybrané výsledky regionálnej prognózy

Kraje	EAO, VZPS, tisíc osôb				Miera nezamestnanosti			
	2000	2007	2013	2017	2000	2007	2013	2017
Bratislavský	335,20	328,74	322,58	311,47	7,24%	3,62%	2,71%	2,82%
Trnavský	269,70	288,96	291,13	286,40	16,37%	6,98%	3,60%	3,24%
Trenčiansky	292,00	294,67	301,01	299,12	15,03%	5,63%	3,27%	3,58%
Nitriansky	333,10	349,50	354,87	350,80	20,82%	13,55%	6,58%	4,76%
Žilinský	328,10	335,81	347,03	347,20	18,49%	10,98%	5,78%	4,73%
Banskobystrický	315,30	329,10	336,62	334,56	21,93%	17,56%	10,70%	8,57%
Prešovský	358,50	379,70	396,83	399,89	22,10%	16,40%	10,42%	8,65%
Košický	355,30	358,52	373,96	376,42	25,57%	18,40%	10,34%	6,77%
SR spolu	2586,90	2665,00	2724,04	2705,87	18,61%	11,97%	6,92%	5,57%
	Nezamestnanosť, VZPS, tis. osôb				zamestnanosť, VZPS, tis. osôb			
	2000	2007	2013	2017	2000	2007	2013	2017
Bratislavský	24,4	11,91	8,75	8,79	310,8	316,8	313,8	302,7
Trnavský	44,7	20,18	10,48	9,29	225,0	268,8	280,6	277,1
Trenčiansky	44,2	16,58	9,85	10,71	247,8	278,1	291,2	288,4
Nitriansky	69,8	47,34	23,34	16,70	263,3	302,2	331,5	334,1
Žilinský	61,4	36,86	20,06	16,42	266,7	298,9	327,0	330,8
Banskobystrický	69,6	57,81	36,02	28,66	245,7	271,3	300,6	305,9
Prešovský	80,1	62,26	41,36	34,59	278,4	317,4	355,5	365,3
Košický	91,2	65,96	38,68	25,47	264,1	292,6	335,3	351,0
SR spolu	485,2	318,9	188,5	150,6	2101,8	2346,1	2535,5	2554,3
	Priemerná hrubá mesačná nominálna mzda, Sk				HDP, s.c.2000, mil. Sk			
	2000	2007	2013	2017	2000	2007	2013	2017
Bratislavský	16072	30480	46138	58105	234762	371107	589471	758341
Trnavský	11500	21863	33088	41446	101207	143941	221085	276401
Trenčiansky	11158	19518	30131	38195	97536	136156	210041	264791
Nitriansky	10598	18482	27410	34626	108812	147975	222726	283161
Žilinský	10707	20405	30010	37382	97422	150425	231762	291575
Banskobystrický	10627	19046	28471	35757	94900	143245	221414	279476
Prešovský	9806	16759	24260	30225	82016	117093	176152	225169
Košický	11908	22150	33286	41364	117819	185609	292799	374979
SR spolu	11799	21735	32324	40633	934473	1395553	2165451	2753893

Použitá literatúra

VAŇO B. 2004, Prognóza vývoja obyvateľov v okresoch SR do roku 2025, Bratislava, Infostat, 2004

KVETAN, V.; MLÝNEK, M.; PÁLENÍK, V.; RADVANSKÝ, M.: Prognóza vývoja ekonomiky SR na roky 2007-2015 s dôrazom na vývoj HDP a zamestnanosti in Pohľady na ekonomiku Slovenska, Slovenská štatistická a demografická spoločnosť, Bratislava, 2007

KVETAN, V.; MLÝNEK, M.; MORVAY, K.; RADVANSKÝ, M.: Aktualizovaná prognóza vývoja sociálno-ekonomickej situácie vo vybraných okresoch SR do roku 2013 (prípady Lučenec a Bratislava), Expertízne štúdie EÚ SAV č. 2. Projekt EQUAL, č. 15110100006, Bratislava, 2006

**Analýza vývoja slovenskej ekonomiky v roku 2006
a prognóza jej vývoja v roku 2007
(s výhľadom do roku 2010)**

Ján Haluška - Michal Olexa
INFOSTAT Bratislava

ABSTRACT:

During two recent years the SR belonged among countries with the highest dynamics of economic development within the EU. In 2006 the real GDP increased by 8.3%, what was by 2 percentage points more than in 2005. Slovakia thereby recorded the fastest economic growth in its history. As compared to the EU average in the previous year, the threefold economic growth was in question.

The authors analyse recent trends in the main fields of the economic development and try to outline the forecast for this year with the extension till 2010.

Citát (ne)známeho prognostika:

„Prognózovanie sociálno-ekonomického, makroekonomického, demografického, ... vývoja je nepretržitý proces. Len čo ukončíte prácu na konkrétnej prognóze, už je neaktuálna a mali by ste začať znova.“

V predloženej štúdii analyzujeme vývojové tendencie vybraných makroekonomických ukazovateľov slovenskej ekonomiky v roku 2006 a na začiatku roka 2007. Zároveň, na základe informácií týkajúcich sa očakávaného vývoja v jej vnútornom a relevantnom vonkajšom prostredí, prognózujeme očakávaný vývoj daných ukazovateľov v roku 2007 a výhľad ich vývoja do roku 2010 (tab. 1 a 2).

Hrubý domáci produkt

Trend zrýchľovania dynamiky hospodárskeho rastu, ktorý už od začiatku minulého roka signalizujú zvyšujúce sa prírastky reálneho HDP, pokračoval aj v závere minulého roka. Tvorba reálneho HDP sa vo 4. štvrťroku 2006 medziročne zvýšila o 9.6%, čo je porovnateľný rast ako vo vlaňajšom v 3. štvrťroku, v ktorom slovenská ekonomika zaznamenala rekordne vysoký prírastok svojej výkonnosti (o 9.8%). Celoročne sa tým jej výkonnosť zvýšila o

8.3%¹, čo na medziročnej báze znamená zrýchlenie hospodárskeho rastu o 2.3 percentuálneho bodu.

Minuloročný rast výkonnosti hospodárstva bol najrýchlejší v doterajšej histórii Slovenska. V súlade s očakávaním k nemu prispel rast domáceho aj vonkajšieho dopytu (celkového vývozu). Kým prvý z týchto agregátov stúpol oproti roku 2005 o 6.4%, druhý až o 20.7%. V porovnaní s rokom 2005 sa však rast domáceho dopytu spomalil o vyše 2 percentuálne body, pretože spotrebiteľský a investičný dopyt rástli pomalšie (stúpili o 6.3% resp. o 7.3%). Jedine konečná spotreba verejnej správy rástla rýchlejšie ako v roku 2005, čo zrejme súvisí so skutočnosťou, že minulý rok bol volebný. V 1. polroku 2006 sa totiž medziročne zvýšila až o 7.2%, v 2. polroku o 1.8%. Napriek tomu bola najpomalšie rastúcou zložkou domáceho dopytu.

Vplyvom rastu vonkajšieho dopytu bol vlani prírastok celkového vývozu takmer o 3 percentuálne body vyšší ako prírastok celkového dovozu. V súvislosti s tým došlo k jednej zásadnej zmene, ktorá je kľúčová z hľadiska dlhodobej udržateľnosti rastu slovenskej ekonomiky a týka sa vplyvu čistého vývozu na vývoj reálneho HDP. Prvý raz v histórii samostatnej SR došlo totiž k tomu, že čistý vývoz prispel pomerne významnou mierou k rastu reálneho HDP, a to v podmienkach vysokého rastu domáceho dopytu.

Jedná sa o kvalitatívne novú podobu rastu slovenskej ekonomiky. V predchádzajúcich rokoch mal čistý vývoz spravidla brzdiaci, výnimočne neutrálny vplyv na rast HDP. V rokoch 1999 resp. 2003 k rastu reálneho HDP tiež významne prispel, bolo to však v podmienkach reálneho poklesu domáceho dopytu (medziročne o 5.9% resp. o 1.3%).

Slovensko patrí v rámci EÚ medzi krajiny s najvyššou dynamikou ekonomického rastu - v porovnaní s priemerom v EÚ bol ekonomický rast v SR vlani až 3 krát vyšší. Avšak úrovňou ekonomického rozvoja oproti krajinám pôvodnej EÚ-15 (pred jej rozšírením v roku 2004) výrazne zaostáva. HDP na obyvateľa v SR (v PPS) dosahoval v roku 2003 podľa údajov Eurostatu 48.3% priemeru v EÚ-15, voči priemeru v EÚ-25 (bez Bulharska a Rumunska) to bolo 52.8%. Zásluhou vysokého ekonomického rastu stúpol HDP na obyvateľa v SR v roku 2005 na 57.1% priemeru EÚ-25 a v roku 2006 na 60.3% (odhad).

Regionálne však existujú v úrovni ekonomického rozvoja na Slovensku výrazné rozdiely. Kým v okolí hlavného mesta Bratislavy dosahovala v roku 2004 úroveň HDP na obyvateľa až 129.3% priemeru EÚ-27 (vrátane Bulharska a Rumunska), západné Slovensko dosiahlo len 52.7%, stredné Slovensko 46.7% a východné Slovensko iba 42.3%.

¹ Kým v 1. polroku 2006 sa výkonnosť slovenskej ekonomiky (meraná prírastkom reálneho HDP) zvýšila o 6.7%, v 2. polroku 2006 o 9.7%.

Aj keď dynamika rastu slovenskej ekonomiky bola v 3. a 4. štvrtroku 2006 takmer identická, dôvody, ktoré k tomu viedli boli identické iba sčasti a súviseli len s vplyvom domáceho efektívneho dopytu ako jedného zo zdrojov ekonomického rastu. Domáci efektívny dopyt (bez vplyvu zmeny stavu zásob) sa totiž v obidvoch týchto štvrtrokoch zvýšil medziročne úplne rovnako - o 5.5%. Avšak dôvody, prečo bol skutočný prírastok reálneho HDP v 3. a 4. štvrtroku oproti prírastku domáceho efektívneho dopytu ešte o vyše 4 percentuálne body väčší, boli rozdielne. Kým v 3. štvrtroku sa o tento dodatočný rastový impulz "postaral" enormný prírastok stavu zásob, vo 4. štvrtroku to bol dôsledok vplyvu čistého vývozu. Čistý vývoz v stálych cenách zaznamenal totiž vo 4. štvrtroku 2006 podstatne lepší výsledok ako vo 4. štvrtroku 2005 (-16.4 mld Sk vs. -32.5 mld Sk) a tým rast reálneho HDP zrýchlil.

Štruktúra rastu HDP, ktorá vznikla vo 4. štvrtroku 2006 sa tohto roku zrejme podstatnejšie nezmení. Hlavným zdrojom ďalšieho zvýšenia výkonnosti slovenskej ekonomiky by teda mal zostať rast domáceho (efektívneho) dopytu, ktorý by mal byť aj tohto roku generovaný najmä investičným dopytom. Relatívne vysoký rast však pravdepodobne dosiahne aj spotrebiteľský dopyt, k čomu na makroúrovni vytvára predpoklady očakávaný rast reálnej mzdy v súvislosti s poklesom inflácie, ako aj rast zamestnanosti a pokles nezamestnanosti. Popri zvyšujúcom sa domácom dopyte by mal rast reálneho HDP podporiť aj čistý vývoz, pričom príspevok čistého vývozu k rastu reálneho HDP bude pravdepodobne vyšší ako vlani. Na základe uvedených predpokladov odhadujeme, že sa tvorba reálneho HDP tohto roku medziročne zvýši o 9.3%. Nominálne by mal byť HDP medziročne vyšší o 11.6% resp. o 190 mld Sk a tým dosiahnuť okolo 1830 mld Sk. Relatívne vysoký ekonomický rast by mal pokračovať aj v ďalších rokoch (tab. 1).

Zahraničný obchod - saldo obchodnej bilancie

Výsledok obchodnej bilancie v roku 2006 bol omnoho horší než sa pred rokom všeobecne očakávalo. Podľa predbežných odhadov stúpol totiž jej deficit oproti roku 2005 o vyše 15 mld Sk a dosiahol cca 93 mld Sk, čo je ekvivalentné 5.6% HDP v bežných cenách. Oproti deficitu v roku 2004 to dokonca predstavuje nárast na takmer dvojnásobok.

Výsledný schodok zahraničného obchodu sa tvoril počas minulého roka viac-menej rovnomerne, ale najväčšia nerovnováha medzi vývozom a dovozom tovarov v bežných cenách vznikla v decembri, keď schodok bilancie dosiahol vyše 12 mld Sk. Prehĺbenie vonkajšej nerovnováhy však bolo v rozhodujúcej miere výsledkom predstihu medziročného rastu dovozu pred rastom vývozu počas prvých troch štvrtrokov minulého roka, ktorý činil v prie-

mere cca 2 percentuálne body. V každom z nich sa totiž deficit obchodnej bilancie medziročne prehĺbil v priemere zhruba o 8 mld Sk. Očakávaný obrat nastal až vo 4. štvrťroku, v ktorom bol schodok obchodnej bilancie medziročne nižší o vyše 10 mld Sk, pretože v poslednom kvartáli rast vývozu naopak predstihol rast dovozu, a to o viac ako 5 percentuálnych bodov. V dôsledku toho bol vlani aj celoročne rast vývozu na medziročnej báze čiastočne vyšší ako rast dovozu (o 24.2% resp. o 23.9%).

Medziročný pokles deficitu obchodnej bilancie vo 4. štvrťroku 2006 sa všeobecne interpretuje ako začiatok nastupujúceho trendu postupného znižovania nerovnováhy medzi vývozom a dovozom tovarov, ktorý by mal v rozhodujúcej miere súvisieť s realizáciou produkcie nových vývozne orientovaných kapacít (Volkswagen, PSA Peugeot a KIA).

Za prvé tri mesiace roka 2007 sa potvrdili očakávania nástupu nového trendu v zahraničnom obchode, ktorým je výrazné zmiernenie nerovnováhy obchodnej bilancie. Údaje za prvý štvrťrok hovoria o zlepšení oproti prvému štvrťroku minulého roka o vyše 20 mld. Sk. Rast je predovšetkým výsledkom značného predstihu vývozu pred dovozom. Dovozy vzrástol oproti roku 2006 o 16 %, rast vývozu bol vyše 25 %.

Deficit obchodnej bilancie za január až marec dosiahol podľa predbežných údajov len 2.7 mld Sk (0.9% HDP). O podstatné zlepšenie vývoja obchodnej bilancie v 1. štvrťroku sa zaslúžil predovšetkým jej prebytok dosiahnutý v januári (cca 5 mld Sk) a výrazný medziročný pokles jej deficitu vo februári. Marcový deficit obchodnej bilancie bol približne na rovnakej úrovni ako v marci minulého.

Rýchly rast vývozu je výsledkom realizácie nových produkčných kapacít v ekonomike, ktorých odbyt je určený predovšetkým pre zahraničné trhy. Podľa podrobných údajov za január a február 2007 sa o to zaslúžili rovnako automobilový aj elektrotechnický priemysel². Z údajov o dovoze je zrejmé výrazné medziročné zníženie dovozu nerastných surovín. Nakoľko ceny ropy od januára stúpali, nižší dovoz môže byť kombináciou nižších cien ropy koncom roka 2006 (kľzavé priemery) a čerpaním zásob surovín s ohľadom na veľmi teplú zimu (dovoz za prvé 2 mesiace bol nižší oproti rovnakému obdobiu 2006 približne o 10 mld. Sk). Predbežné údaje o zahraničnom obchode od začiatku roka rovnako nenaznačujú, že by sa bolo uskutočnilo očakávané predzásobovanie automobilových koncernov, vzhľadom na plánované spúšťanie plnej výroby dopravných zariadení (PSA, KIA).

² Medziročný priemerný nárast vývozu Strojov, prístrojov a el. zariadení (16. trieda tovarov podľa harmonizovaného systému) za 2 mesiace bol na úrovni 33 %, vývoz v 17. triede (Lietadlá, vozidlá, plavidlá a dopravné zariadenia) vzrástol medziročne dvojnásobne.

Export by si mal udržať vysoké tempa rastu do konca roka, nakoľko sa očakáva smerovanie k využívaniu plných kapacít výroby automobilového i elektrotechnického priemyslu. Pri dovoze očakávame vyššie nominálne dovozy nerastných surovín a ich tesnejšie približovanie sa k mesačným hodnotám minulého roku. Ďalším faktorom podporujúcim vyšší rastu dovozu sú očakávané veľké investície (Samsung) a predzásobovanie sa automobilového priemyslu na plnú výrobu. Aj napriek týmto predpokladom by sa mal deficit obchodnej bilancie za rok 2007 dostať na výrazne nižšie úrovne oproti 2006. Oproti minulému roku by mal byť podstatne nižší a dosiahnuť úroveň 25 - 30 mld. Sk.

Spotrebiteľské ceny

V porovnaní s rokom 2005 bola dynamika rastu cien na spotrebiteľskom trhu v minulom roku vyššia. Úhrnná hladina spotrebiteľských cien stúpila o 4.5% (v priemere), čo je prírastok o 1.8 percentuálneho bodu vyšší ako v roku 2005. Aj keď nárast medziročnej inflácie súvisel predovšetkým so zvýšením viacerých regulovaných cien v januári resp. v novembri minulého roka (plyn), čiastočne ide aj na vrub zvýšenia jadrovej inflácie. Jej relatívne mierny nárast - z 1.1% v roku 2005 na 2.5% v roku 2006 - bol totiž do značnej miery dôsledkom obratu vo vývoji cien potravín, ktoré po prechodnom medziročnom poklese v roku 2005 (o 1.4%) vlni medziročne o rovnaké percento stúpili.

Pokles medziročnej miery inflácie, ktorý začal v septembri 2006, v priebehu 1. štvrtroka 2007 - v súlade s očakávaním - pokračoval. Súviselo to najmä so skutočnosťou, že regulované ceny energií sa v januári tohto roka nielenže nemenili (na rozdiel od minulých rokov), ale vodné a stočné, ako aj zemný plyn pre domácnosti zlacneli. Spotrebiteľské ceny preto na začiatku tohto roka zaznamenali netradične nízky medzimesačný prírastok - v januári 2007 boli oproti decembru 2006 vyššie len o 1%. Vzhľadom na to, že vo februári stúpili len o 0.2% a v marci sa ich rast zastavil, kumulatívne sa v 1. štvrtroku zvýšili oproti decembru 2006 iba o 1.2%. V rovnakom období vlni bol ich kumulatívny nárast viac ako dvojnásobne vyšší (oproti decembru 2005 stúpili o 2.7%).

Vplyvom spomaľujúcej sa dynamiky rastu spotrebiteľských cien sa medziročná miera celkovej inflácie znížila zo 4.2% v decembri 2006 na 3.0% v januári 2007, resp. na 2.7% vo februári a v marci. Na medziročnej báze teda úhrnná hladina spotrebiteľských cien stúpila v 1. štvrtroku 2007 v priemere o 2.8%, čo je o 1.5 percentuálneho bodu menej ako v rovnakom období minulého roka. K uvedenému poklesu inflácie prispeli relatívne najviac ceny v doprave a ceny služieb pôšt a telekomunikácií, keďže boli medziročne nižšie v priemere o 3.5%,

resp. o 1.1%. Ceny všetkých ostatných položiek v rámci štruktúry COICOP v 1. štvrtroku stúpili³.

V apríli sa rast úhrnnej hladiny spotrebiteľských cien síce obnovil, ale jej medzimesačný prírastok bol relatívne nízky. Oproti marcu sa tieto ceny totiž zdvihli len o 0.2%, k čomu prispeli relatívne najviac ceny v doprave a ceny potravín, ktoré stúpili rovnako o 0.7%. Nadpriemerný prírastok však dosiahli aj ceny odevov a obuvi (o 0.3%), ale - vzhľadom na pokles DPH na ceny liekov (z 19% na 10%) - zaznamenali relatívne vysoký prírastok tiež ceny v zdravotníctve (o 0.2%). Hoci sa úhrnná hladina cien na spotrebiteľskom trhu v apríli zvýšila, medziročná miera celkovej inflácie sa nezmenila a zostala na februárovej, resp. marcovej úrovni, teda na 2.7%⁴.

Ceny v doprave a ceny potravín treba považovať za najväčšie riziká aj z hľadiska odhadu ďalšieho vývoja cien na spotrebiteľskom trhu v tomto roku. Predpokladáme však, že žiaden z uvedených rizikových faktorov nespôsobí v horizonte tejto predikcie výraznejší inflačný impulz. Keďže vývoj výmenného kurzu koruny by mal aj naďalej generovať tlak na pokles úhrnnej hladiny spotrebiteľských cien, možno s veľkou pravdepodobnosťou očakávať, že doterajší trend postupného, ale veľmi mierneho poklesu medziročnej miery inflácie bude pokračovať. Dočasne však môže byť prerušený vplyvom sezónnosti vývoja cien potravín v jesenných mesiacoch. Na základe uvedených predpokladov odhadujeme, že v decembri tohto roku bude medziročná miera inflácie predstavovať 2.4%, pričom priemerná miera inflácie by mala v tomto roku dosiahnuť 2.6%.

Nominálne a reálne mzdy

Rast priemernej mesačnej nominálnej mzdy sa vlani vo väčšine odvetví, v ktorých sa jej vývoj sleduje pravidelne aj s mesačnou periodicitou, spomalil⁵. Opačný vývoj zaznamenali len dve odvetvia, menovite "Maloobchod" a "Predaj a údržba vozidiel". V nich sa jej rast oproti roku 2005 zrýchlil a podľa predbežných údajov dosiahol okolo 10%, čo z hľadiska vývoja priemernej mesačnej nominálnej mzdy v celom hospodárstve v roku 2006 predstavuje

³ Spomedzi nich sa relatívne najviac zvýšili ceny bývania, vody, elektriny a plynu (medziročne v priemere o 6.2%). Z hľadiska úhrnu však zaznamenali nadpriemerný prírastok ešte aj ceny alkoholických nápojov a tabaku (o 5.1%), ceny vzdelávania (o 4.0%), ceny v zdravotníctve (o 3.6%), ale aj ceny potravín a nealkoholických nápojov (o 3.2%).

⁴ Takto stabilizovaná miera medziročnej inflácie je podmienená aj vývojom jadrovej a čistej inflácie, ktorý nepoukazuje - napriek vysokému tempu rastu ekonomiky v 1. štvrtroku - na prítomnosť dopytových tlakov. NBS v dôsledku toho znížila kľúčovú sadzbu v marci zo 4.75% na 4.50% a v apríli na 4.25%.

⁵ Jedná sa o 9 vybraných odvetví, ktorých podiel na celkovej zamestnanosti dosiahol v minulom roku takmer 62%.

nadpriemerný rast. Na druhej strane, v priemysle (najmä spracovateľskom) a stavebníctve bol vlni rast nominálnej mzdy podpriemerný (o 5.8% resp. o 5.0%), zatiaľ čo v odvetví telekomunikácií bol veľmi blízky k rastu nominálnej mzdy v celom hospodárstve. V každom z deviatich odvetví však medziročný relatívny prírastok nominálnej mzdy prevýšil úroveň minulo-ročnej (priemernej) inflácie, z čoho vyplýva, že v každom z nich došlo aj k rastu reálnej mzdy - v rozpätí od 0.5% (v stavebníctve) do 5.9% (v maloobchode).

Jednoznačne najvyššie absolútne prírastky nominálnej mzdy dosahuje už niekoľko rokov odvetvie telekomunikácií. Ani minulý rok nebol z tohto hľadiska výnimkou. Relatívny medziročný prírastok nominálnej mzdy v tomto odvetví o 7.8% (dosahujúci zhruba úroveň priemerného prírastku nominálnej mzdy v celom hospodárstve) totiž v absolútnej podobe znamená prírastok takmer o 3 000 Sk, čo je najviac spomedzi všetkých deviatich odvetví. Nominálna mzda v odvetví telekomunikácií tým vlni stúpila takmer na 41 tisíc Sk. V porovnaní s priemernou nominálnou mzdou v celej ekonomike resp. v priemysle to predstavuje viac ako dvojnásobok. Zároveň je zhruba trojnásobkom priemernej nominálnej mzdy v stavebníctve resp. maloobchode a štvornásobkom priemernej nominálnej mzdy v hoteloch a reštauráciách.

K spomaleniu rastu priemernej mesačnej nominálnej mzdy došlo vlni aj v celom hospodárstve - z 9.2% v roku 2005 na 8.0%. Medziročne sa jej úroveň zvýšila takmer o 1 500 Sk a dosiahla 18 761 Sk. Rast reálnej mzdy sa spomalil ešte výraznejšie - zo 6.3% v roku 2005 na 3.3% v roku 2006 - pretože pomalší rast nominálnej mzdy bol vlni sprevádzaný vyššou infláciou. Na druhej strane sa tým však obnovila priaznivá makroekonomická relácia. Vlni totiž rast súhrnnej nominálnej resp. reálnej produktivity práce (o 8.8% resp. o 5.9%) predstihol na makroúrovni rast nominálnej resp. reálnej mzdy (na rozdiel od roku 2005).

Vplyvom očakávaného vysokého rastu súhrnnej produktivity práce by sa mal relatívne vysoký medziročný rast nominálnej mzdy v ďalších štvrtrokoch tohto roka nielen udržať, ale aj čiastočne zrýchliť. Potrebné predpoklady pre to vytvára existujúci predstih rastu produktivity pred rastom miezd. Na základe toho odhadujeme, že priemerná mesačná nominálna mzda na makroúrovni sa tohto roku medziročne zvýši o 7.3%, v dôsledku čoho prekročí hranicu 20 000 Sk. Vzhľadom na očakávanú úroveň priemernej inflácie v tomto roku (2.6%) to znamená rast reálnej mzdy o 4.6%, čo oproti roku 2006 predstavuje zrýchlenie jej rastu o 1.3 percentuálneho bodu.

Trh práce - zamestnanosť a nezamestnanosť

V porovnaní s predchádzajúcim obdobím bol najvýraznejšou zmenou vývoja odvetvovej zamestnanosti v minulom roku pokles zamestnanosti v priemysle. Ide o pomerne prekvapivý obrat, pretože zamestnanosť v tomto odvetví predtým počas piatich rokov sústavne rástla. Podľa predbežných údajov sa však vlani jej úroveň znížila o 1.3%, čo predstavuje úbytok cca 10 tisíc zamestnancov (oproti roku 2005). Rozhodujúca časť z nich bola uvoľnená z priemyselnej výroby (cca 6 tisíc), k poklesu zamestnanosti však došlo aj vo výrobe a rozvoде elektriny, plynu a vody, ako aj v ťažbe nerastných surovín.

V ostatných odvetviach vývoj zamestnanosti v minulom roku spravidla zvýraznil pozitívne trendy z predchádzajúcich rokov. Platí to predovšetkým o stavebníctve, ktoré dosiahlo opäť výrazné zrýchlenie rastu zamestnanosti (zo 6.9% v roku 2005 na 9.3% v roku 2006), ale veľmi vysoký rast zamestnanosti pokračoval aj vo veľkoobchode (o 9.2%), v hoteloch a reštauráciách (o 8.8%), v predaji a údržbe vozidiel (o 8.4%) a v maloobchode (o 7.4%). Navyše sa vlani zastavil dlhodobý pokles zamestnanosti v odvetví "Doprava a skladovanie" (trval od roku 2000) a výrazne sa spomalil trend poklesu zamestnanosti v odvetví "Pošty a telekomunikácie" (začal v roku 2002). Vo všetkých 9 odvetviach, v ktorých sa vývoj zamestnanosti sleduje pravidelne aj s mesačnou periodicitou, sa vlani počet zamestnancov zvýšil o 2.9% (v priemere).

Vysoký prírastok odvetvovej zamestnanosti sa odzrkadlil aj na makroúrovni. Zo štvrtročného štatistického výkazníctva (vrátane živnostníkov) totiž vyplýva rast celovej zamestnanosti o 2.2%, čo predstavuje vyše 70 tisíc nových pracovných miest. Podľa výberových zisťovaní pracovných síl (VZPS) bol rast počtu pracujúcich v roku 2006 dokonca ešte vyšší - o 3.8%⁶, v súvislosti s čím výrazne klesol počet nezamestnaných - v priemere o vyše 74 tisíc osôb resp. o 17.3% oproti roku 2005. Miera nezamestnanosti z VZPS sa znížila na 13.3% (zo 16.2% v roku 2005), ale v poslednom štvrtroku minulého roka klesla až na 12.0%, čo v absolútnom vyjadrení predstavuje 319 tisíc nezamestnaných osôb (medziročný úbytok o vyše 88 tisíc osôb).

Trh práce v SR je - vzhľadom na počet obyvateľov - malým trhom, pričom veľkosťou je v rámci krajín EÚ porovnateľný s trhom práce vo Fínsku. Počet ekonomicky aktívneho obyvateľstva sa v posledných šiestich rokoch výraznejšie nezmenil (osciluje okolo 2650 tisíc osôb), predtým však mal rastúci trend.

⁶ Tento údaj však zahŕňa aj tých pracujúcich, ktorí si vlani našli prácu v zahraničí. Ak by sme vplyv tohto počtu vylúčili, potom by prírastok zamestnanosti predstavoval 2.3%.

Miera ekonomickej aktivity sa dlhodobo udržiava v blízkosti hranice 60%⁷. Aj miera zamestnanosti dosahovala až do roku 1998 cca 60% (z toho 67% u mužov a 54% u žien), ale vplyvom poklesu zamestnanosti v rokoch 1999 a 2000 (kumulatívne o vyše 4%) sa znížila na zhruba 57% (u mužov na cca 63%, u žien na 51%). Odvtedy sa v dôsledku obnovenia rastu zamestnanosti opäť zvyšuje. V roku 2006 stúpila na vyše 59%, pričom vo 4. štvrťroku 2006 dosiahla takmer 60%.

V roku 2005 bola miera zamestnanosti v SR porovnateľná s mierou zamestnanosti v Maďarsku resp. v Taliansku, v porovnaní s mierou zamestnanosti v EÚ-25 však bola nižšia cca o 6 percentuálnych bodov.

Sprievodným javom vývoja slovenskej ekonomiky, ktorý však vznikol ešte v čase keď bola súčasťou bývalej ČSFR, je vysoká miera nezamestnanosti. Jej rast kulminoval v 1. štvrťroku 2001, v ktorom stúpila až na 19.7%. Odvtedy sa miera nezamestnanosti postupne znižuje, avšak predpoklady pre jej trvalejší pokles vytvorili až zásadné ekonomické reformy, ktoré boli realizované od roku 2002. V súvislosti so zvyšujúcim sa prílevom PZI, ktorý generuje vznik nových pracovných miest, sa miera nezamestnanosti znížila v roku 2006 na 13.3% (v priemere). Vo 4. štvrťroku 2006 však klesla až na 12.0%, teda na najnižšiu úroveň od roku 1997. Naďalej však existujú výrazné regionálne rozdiely v miere nezamestnanosti. Kým v okolí hlavného mesta Bratislavy nedosahovala miera evidovanej nezamestnanosti na konci minulého roka ani 3%, v juhovýchodnej časti SR existujú regióny, v ktorých prevyšovala 25%.

Špecifickým problémom trhu práce v SR, ktorý sa navyše stále prehĺbuje, je dlhodobá nezamestnanosť. Podiel dlhodobo nezamestnaných na celkovom počte nezamestnaných bol v SR vysoký už na začiatku jej samostatnej existencie (1993), keď dosahoval okolo 40%. Odvtedy sa však sústavne zvyšuje, pričom rapídne stúpol najmä v posledných dvoch rokoch - v roku 2005 na vyše 68% a v roku 2006 už prekročil 73%.

V rámci EÚ-25 má SR ďaleko najvyšší podiel dlhodobo nezamestnaných, značne prevyšujúci priemer v EÚ-25, ktorý v roku 2005 predstavoval 45.0%.

Očakávaný vysoký ekonomický rast vytvára predpoklady na to, aby pozitívne tendencie vývoja na trhu práce pokračovali aj v roku 2007 a s menšou intenzitou aj v ďalšom období. Vplyvom rastu dopytu na trhu práce by teda mal pokračovať tak rast celkovej zamestnanosti, ako aj pokles nezamestnanosti (tab. 2).

⁷ V SR tvorili do roku 2003 obyvateľstvo v produktívnom veku muži vo veku 15-59 rokov a ženy vo veku 15-54 rokov. V dôsledku prebiehajúcej dôchodkovej reformy sa vek odchodu do dôchodku postupne zvyšuje na 62 rokov jednotne pre mužov aj ženy.

Makroekonomický rámec očakávaného vývoja slovenskej ekonomiky v tomto roku vytvára základné predpoklady pre udržateľnosť aktuálnych vývojových tendencií na trhu práce. Dôsledkom pokračujúceho rastu dopytu na tomto trhu by mal byť celoročne prírastok celkovej zamestnanosti, ktorý bude porovnateľný s jej prírastkami v minulých dvoch rokoch. Kým v metodike štatistických zisťovaní by mal jej medziročný prírastok dosiahnuť 2.4%, v metodike VZPS až 2.9%.

V súvislosti s tým by mal pokračovať aj trend zmierňovania nerovnováhy na trhu práce. Počet nezamestnaných podľa VZPS by mal už v 2. štvrťroku klesnúť pod hranicu 300 tisíc osôb, pričom ich priemerný počet v tomto roku by mal predstavovať 291 tisíc osôb. Oproti minulému roku by to znamenalo úbytok cca o 60 tisíc osôb (v priemere). Tým by sa v 2. polroku vytvorili predpoklady pre pokles miery nezamestnanosti pod hranicu 11%, pričom tesne pod túto hranicu by mala zároveň dostať aj priemerná miera nezamestnanosti v tomto roku.

Ing. Ján Haluška, PhD. – haluska@infostat.sk

Ing. Michal Olexa, PhD. – olexa@infostat.sk

INFOSTAT

Dúbravská cesta 3

845 24 Bratislava 45

	2003	%chg	2004	%chg	2005	%chg	2006	%chg	2007	%chg	2008	%chg	2009	%chg	2010	%chg
VYBRANÉ MAKROEKONOMICKÉ UKAZOVATELE																
Deflátor HDP (1995=1) (PY)	1,151	4,7	1,220	6,0	1,249	2,4	1,283	2,7	1,311	2,1	1,334	1,8	1,357	1,7	1,383	1,9
Index spotrebiteľských cien priemer, (dec.2000=1),(CPI)	1,181	8,5	1,270	7,5	1,305	2,7	1,363	4,5	1,399	2,6	1,425	1,9	1,451	1,8	1,476	1,8
Priemerná nominálna mesačná mzda, Sk (W)	14 438	6,3	15 905	10,2	17 364	9,2	18 761	8,0	20 131	7,3	21 479	6,7	22 811	6,2	24 157	5,9
Priemerná reálna mesačná mzda, SKK (W/CPI)	12 222	-2,1	12 520	2,4	13 307	6,3	13 762	3,4	14 395	4,6	15 071	4,7	15 734	4,4	16 364	4,0
Priemerný počet zamestnaných osôb v NH spolu, mil. osôb (L)	2,051	0,8	2,056	0,3	2,101	2,2	2,148	2,2	2,200	2,4	2,222	1,0	2,240	0,8	2,257	0,8
Počet pracujúcich, mil. osôb, priemer (L_ILO)	2,165	1,8	2,170	0,3	2,216	2,1	2,301	3,8	2,368	2,9	2,397	1,2	2,421	1,0	2,445	1,0
Počet nezamestnaných mil. osôb, priemer (LU_ILO)	0,459	-5,7	0,481	4,7	0,427	-11,1	0,353	-17,3	0,291	17,6	0,278	-4,5	0,264	-5,0	0,256	-3,2
Miera nezamestnanosti v %, (RU_ILO)	17,4	*	18,1	*	16,2	*	13,3	*	10,9	*	10,7	*	10,3	*	9,9	*
Reálna produktivita práce (Y/L)	513,9	3,3	540,3	5,1	560,6	3,8	593,6	5,9	633,68	6,7	670,1	5,7	703,3	5,0	736,1	4,7
Nominálna produktivita práce (YP/L)	591,4	8,2	659,2	11,5	700,2	6,2	761,7	8,8	830,5	9,0	893,98	7,6	954,3	6,7	1017,7	6,7

**RÝCHLE ODHADY VÝVOJA
ZÁKLADNÝCH MAKROEKONOMICKÝCH UKAZOVATEĽOV
SLOVENSKEJ EKONOMIKY
(Metodológia a výsledky)**

*Ján Haluška
INFOSTAT Bratislava*

Abstract: In INFOSTAT, the activities in building and verification of a specific model framework supporting the preparation of flash estimates started in 2004. Using quarterly time series, different types of econometric models for real GDP and total employment were created and estimated. Models are based on ECM methodology.

Key words: GDP, total employment, business tendency surveys, quantitative and qualitative indicators, time series, seasonality, stationarity, error correction model

1. Úvod

Zostavovanie rýchlych a zároveň objektivizovaných odhadov vývoja základných makroekonomických ukazovateľov je vo vyspelom svete založené na využívaní modelových nástrojov. Takáto prax zostavovania rýchlych odhadov existuje už dnes aj v rezorte Štatistického úradu Slovenskej republiky (ŠÚ SR), z objektívnych príčin však ešte nemá dlhú históriu. Predpoklady pre systematické využívanie modelových nástrojov na uvedené účely v podmienkach SR vznikali totiž postupne, jednak v závislosti na rozširovaní štatistickej údajovej základne a zvyšovaní jej kvality, jednak v súvislosti so znižovaním nestability vývoja samotných makroekonomických ukazovateľov v dôsledku postupujúceho procesu ekonomickej transformácie.

V rezorte ŠÚ SR sa výskumom v oblasti konštrukcie a použitia modelových nástrojov dlhodobo zaoberá Inštitút informatiky a štatistiky (INFOSTAT) v Bratislave. Motiváciou pre konštrukciu modelových nástrojov pre zostavovanie rýchlych odhadov bola skutočnosť, že od roku 2005 je ŠÚ SR povinný v predstihu pripravovať a zverejňovať rýchle odhady vývoja HDP a celkovej zamestnanosti, a to vždy do 45 dní po skončení štvrťroka. Je to teda o 15 dní skôr ako zverejňuje komplexnú publikáciu o vývoji slovenskej ekonomiky v príslušnom štvrťroku.

Modelové nástroje, ktoré INFOSTAT v súčasnosti podporne využíva na zostavovanie rýchlych odhadov, sú výsledkom riešenia úlohy "Vývoj modelov pre štatistické

makroekonomické analýzy a prognózy vývoja ekonomiky SR"¹. Z metodologického hľadiska sa jedná o modely ekonometrického typu, ktoré umožňujú zostavovať rýchle odhady vývoja uvedených makroekonomických ukazovateľov na základe informácií vyplývajúcich z časových radov kvantitatívneho charakteru, ako aj z časových radov kvalitatívneho charakteru, teda z výsledkov konjunkturálnych prieskumov.

2. Modelový aparát na rýchle odhady vývoja HDP a zamestnanosti

Modelový aparát na rýchle odhady tvoria jednorovnicové ekonometrické modely zobrazujúce vývoj HDP a celkovej zamestnanosti. Ich konštrukcia a špecifikácia využíva princípy modelov s členom korigujúcim chyby (*Error Correction Models - ECM*), čo znamená, že každý z nich vyjadruje separátne krátkodobý i dlhodobý vplyv jednotlivých vysvetľujúcich premenných na vývoj príslušnej vysvetľovanej premennej. Na kvantifikáciu parametrov boli použité pôvodné štvrťročné časové rady príslušných ukazovateľov (teda sezónne neočistené časové rady) v kombinácii so sezónnymi premennými (filtrami).

Každý z jednorovnicových modelov existuje v troch rôznych verziách, ktoré sa navzájom odlišujú rozdielnou špecifikáciou, t.j. použitými vysvetľujúcimi faktormi. Vo všeobecnosti je totiž možné pristupovať k zostavovaniu (štvrťročných) rýchlych odhadov uvedených ukazovateľov alternatívne, a to

- výlučne len na základe využitia informácií kvalitatívneho charakteru, t.j. na výsledkoch konjunkturálnych prieskumov,
- výlučne len na základe využitia informácií kvantitatívneho charakteru, t.j. pomocou mesačných časových radov relevantných odvetvových ukazovateľov,
- kombináciou využitia oboch typov informácií súčasne.

V prvom prípade je pre rýchly odhad HDP rozhodujúci vývoj indikátora ekonomického sentimentu (IES) alebo jeho jednotlivých zložiek, t.j. indikátora dôvery v priemysle, stavebníctve, maloobchode a indikátora spotrebiteľskej dôvery. Pre rýchly odhad celkovej zamestnanosti je to indikátor očakávanej zamestnanosti v priemysle, stavebníctve a maloobchode. V druhom prípade sú pre rýchly odhad HDP východiskom mesačné časové rady vývoja priemyselnej produkcie, stavebnej produkcie, maloobchodného obratu, vývozu tovarov a dovozu tovarov. Pre rýchly odhad celkovej zamestnanosti predstavujú východisko

¹ Ako úlohu štátnej objednávky ju INFOSTAT riešil v rokoch 2002-2005, pričom v jednotlivých etapách riešenia boli postupne vytvorené a overené viaceré modelové nástroje [1, 2, 3].

mesačné časové rady o vývoji zamestnanosti vo vybraných odvetviach hospodárstva. V treťom prípade sú primárnymi zdrojmi informácií pre rýchly odhad HDP a celkovej zamestnanosti všetky vyššie uvedené štatistické informácie. Ktorá z nich sa nakoniec uplatní, závisí vždy od výsledkov regresnej a korelačnej analýzy.

Metodologický postup konštrukcie a odhadu ECM je vo všetkých troch prípadoch zložený z dvoch krokov. V prvom kroku sa odhadne dlhodobý rovnovážny vzťah medzi nestacionárnymi premennými². Pomocou časového radu rezíduí vyplývajúceho z dlhodobého vzťahu sa v druhom kroku odhadne ekonometrický model v tvare ECM, v ktorom sa odchýlka od dlhohodobej rovnováhy objavujúca sa v jednom období čiastočne koriguje v nasledujúcom období. Všeobecne platí, že ak dlhodobý vzťah má tvar

$$Y_t = a + b * X_t \quad \text{kde } X \text{ môže byť vektor premenných}$$

potom jemu zodpovedajúci ECM má tvar

$$\Delta Y_t = c \Delta X_t - \lambda * (Y_{t-1} - a - b * X_{t-1})$$

$$\text{kde } \Delta Y_t = Y_t - Y_{t-1} \quad \text{a} \quad \Delta X_t = X_t - X_{t-1}$$

Predpokladá sa, že parameter λ je kladný, pričom jeho hodnota sa interpretuje ako časť z "chyby" (t.j. skutočnej odchýlky z dlhodobého vzťahu), ktorá je korigovaná v nasledujúcom období. Získaný model v tvare ECM sa vyznačuje vysokou vnútornou stabilitou, avšak len vtedy, keď sú všetky jeho tri zložky stacionárne. V tom prípade je riziko explozívneho vývoja pri dynamickej simulácii minimálne.

3. Záver

Z doterajších skúseností vyplýva, že viaceré z uvedených ukazovateľov kvalitatívneho i kvantitatívneho charakteru možno z hľadiska ich použitia na rýchle odhady považovať za štatisticky významné, a to tak separátne, ako aj v kombinácii s niektorými inými ukazovateľmi. Zároveň však platí, že presnosť rýchlych odhadov získaných na základe jednotlivých verzií modelového aparátu je rozdielna. Podľa očakávania majú relatívne

² Časové rady makroekonomických ukazovateľov sú spravidla nestacionárne, čo znamená, že ich priemer a rozptyl sa v čase menia. Použitie príslušných štatistických testov však ukazuje, že časové rady prvých diferencií týchto ukazovateľov sú už väčšinou stacionárne.

najvyššiu presnosť tie rýchle odhady, ktoré sú založené na kombinácii využitia oboch typov informácií súčasne. V prílohe sú výsledky rýchlych odhadov medziročných relatívnych zmien reálneho HDP, zložiek jeho použitia a celkovej zamestnanosti za 1.-4. štvrťrok 2006 uvedené spolu s ich skutočnými relatívnymi zmenami.

Naopak, relatívne najnižšiu presnosť vykazujú rýchle odhady, ktoré zohľadňujú len výsledky konjunkturálnych prieskumov. Ukazuje sa, že rýchle odhady HDP získané len na základe vývoja IES sú spravidla podhodnotené. Súvisí to pravdepodobne so skutočnosťou, že IES zatiaľ nepokrýva sektor služieb, ktorého podiel na tvorbe HDP je však v našich podmienkach vysoký, keďže prevyšuje 50%.

4. Literatúra

1. Štvrťročný ekonometrický model reálneho sektora slovenskej ekonomiky. INFOSTAT, november 2003.
2. Modelové nástroje na prognózovanie krátkodobého vývoja vybraných ukazovateľov slovenskej ekonomiky. INFOSTAT, november 2004.
3. Modelový aparát na rýchle odhady vývoja HDP a zamestnanosti (Využitie konjunkturálnych prieskumov a zložených indikátorov). INFOSTAT, november 2005.

Adresa autora:

Ján Haluška, Ing., PhD.
INFOSTAT, Dúbravská cesta 3
845 24 Bratislava 45
haluska@infostat.sk

**Rýchly odhad HDP a celkovej zamestnanosti
v 1.-4. štvrtroku 2006
(medziročné prírastky v %)**

HDP

	<i>Rýchly odhad</i>	<i>Skutočnosť</i>
1. štvrtrok 2006	7.2	6.7
2. štvrtrok 2006	5.7	6.7
3. štvrtrok 2006	6.9	9.8
4. štvrtrok 2006	9.0	9.6

Konečná spotreba domácností

	<i>Rýchly odhad</i>	<i>Skutočnosť</i>
1. štvrtrok 2006	4.3	6.6
2. štvrtrok 2006	5.5	5.9
3. štvrtrok 2006	7.0	6.5
4. štvrtrok 2006	6.9	6.1

Konečná spotreba verejnej správy

	<i>Rýchly odhad</i>	<i>Skutočnosť</i>
1. štvrtrok 2006	4.9	7.8
2. štvrtrok 2006	5.6	6.6
3. štvrtrok 2006	2.3	1.2
4. štvrtrok 2006	4.7	2.3

Tvorba hrubého fixného kapitálu

	<i>Rýchly odhad</i>	<i>Skutočnosť</i>
1. štvrtrok 2006	15.3	13.8
2. štvrtrok 2006	9.1	3.6
3. štvrtrok 2006	7.2	6.7
4. štvrtrok 2006	3.5	7.0

Vývoz výrobkov a služieb

	<i>Rýchly odhad</i>	<i>Skutočnosť</i>
1. štvrtrok 2006	18.5	17.7
2. štvrtrok 2006	14.1	18.1
3. štvrtrok 2006	22.8	23.8
4. štvrtrok 2006	26.7	22.6

Dovoz výrobkov a služieb

	<i>Rýchly odhad</i>	<i>Skutočnosť</i>
1. štvrťrok 2006	22.0	19.6
2. štvrťrok 2006	14.6	14.0
3. štvrťrok 2006	21.9	22.9
4. štvrťrok 2006	17.9	15.3

Celková zamestnanosť

	<i>Rýchly odhad</i>	<i>Skutočnosť</i>
1. štvrťrok 2006	2.1	2.1
2. štvrťrok 2006	2.3	2.3
3. štvrťrok 2006	2.0	2.5
4. štvrťrok 2006	2.1	2.2

Analýza vplyvu úrovni zvoleného faktora na variabilitu hodnôt ekonomických ukazovateľov (ANOVA) – formulácia úlohy, model, postup analýzy, príklad

Doc. Ing. Jozef Chajdiak, CSc. – Statis, Bratislava

Abstract: The paper consists of the analysis by One-Way ANOVA. This is analyzing the effect of levels juridical status FORMA to variance of the Profit on the file of firms produced electrical machine and device in 2005 year.

Máme k dispozícii súbor údajov o firmách. Nech súbor obsahuje údaje o n firmách (rozsah súboru je n). Firmy sú nositeľkami určitých ekonomických vlastností, ktoré nachádzajú svoje konkrétne vyjadrenie v dvoch ukazovateľoch. Prvý ukazovateľ – premenná X – vyjadruje určitú vlastnosť, napríklad V_{64} – objem výsledku hospodárenia po zdanení v tisíc Sk (V označuje Výkaz ziskov a strát Uč POD 2-01 a 64 označuje 64. riadok v tomto výkaze). Vlastností je viacero, sú vyjadrené inými riadkami Výkazu ziskov a strát alebo výkazu Súvaha. V tomto príspevku bude realizovaná analýza len jedného ukazovateľa X (analýza iných ukazovateľov je analogická). Druhý ukazovateľ – premenná A – vyjadruje príslušnosť firmy k určitej klasifikačnej vlastnosti, napríklad premenná FORMA – právna forma firmy znamená, že jej hodnota 112 reprezentuje spoločnosť s ručením obmedzeným, 121 reprezentuje akciovú spoločnosť, ... atď. a 000 je hodnota pre nezistenú právnu formu firmy. Opäť klasifikačných premenných môže byť viacero (klasifikácia podľa OKEČ, podľa územnej príslušnosti, druhu vlastníctva,...) a v tomto príspevku bude v analýze využitá len jedna klasifikačná vlastnosť A .

Postup analýzy

V prvom kroku analýzy sa realizuje základný štatistický rozbor premennej X . Zisťuje sa počet nechýbajúcich hodnôt n , minimálna a maximálna hodnota $x_{\min}$ a $x_{\max}$ v súbore, priemerná hodnota $\bar{x}$, medián $x_{0,50}$ a smerodajná odchýlka s_x . Doplňujúco sa môžu určiť tiež hodnoty celkového rozsahu súboru, počtu chýbajúcich hodnôt, úhrnu hodnôt, miery asymetrie, miery špicatosti, či ďalšie charakteristiky. Súčasťou prvého kroku analýzy je určenie radu rozdelenia početností výskytu hodnôt ukazovateľa X v tvare frekvenčnej tabuľky a histogramu.

Súčasťou druhého kroku analýz môžu byť intervalové odhady štatistík použitých v prvom kroku (tradičným je 95-percentný interval spoľahlivosti pre strednú hodnotu) a testy hypotéz o strednej hodnote, variabilite, úhrne a tvare a parametroch radu rozdelenia početností.

V treťom kroku analýz môžeme analyzovať ďalšie zreteľahodné skutočnosti. Z pohľadu tohto príspevku, v prípade, že zistíme, že variabilita hodnôt premennej X v našom súbore je nenulová, môžeme si klásť otázku „Čo spôsobuje

variabilitu hodnôt premennej X ?“. Odpoveď môže byť pomerne jednoduchá: „Úrovne faktora A !“. Kvantifikáciu tejto odpovede získame použitím riešenia úlohy ANOVA (ANalysis Of VAriance) – analýzy rozptylu.

Postup riešenia úlohy analýzy rozptylu (ANOVA)

Riešenie úlohy analýzy rozptylu obsahuje niekoľko okruhov otázok a odpovedí:

- opisné štatistiky pre jednotlivé úrovne faktora A a súbor spolu,
- grafická prezentácia vybraných štatistík,
- model riešenia úlohy analýzy rozptylu (špecifikácia efektov α_i),
- overenie nevýznamnosti vplyvu úrovní faktora A na variabilitu premennej X ,
- špecifikácia podielu modelom (faktorom A) vysvetlenej variability premennej X (R^2),
- prípadná analýza ďalších aspektov (napríklad analýza homoskedasticity, analýza kontrastov,...),
- praktické závery.

Premenná X nadobúda hodnoty x_i u jednotlivých firiem $i = 1, 2, \dots, n$. Faktor A má m úrovní $a_1, a_2, \dots, a_m$ a u každej firmy nadobúda nejakú konkrétnu jednoznačnú hodnotu.

Opisné štatistiky pre jednotlivé úrovne faktora A a súbor spolu sú usporiadané v nasledujúcej tabuľke:

Tabuľka opisných štatistík

Úroveň	Rozsah	Minimum	Maximum	Priemer	Medián	Std
a_1	n_{a1}	$\min_{a1}$	$\max_{a1}$	$\bar{x}_{a1}$	$x_{a1;0,50}$	s_{a1}
a_2	n_{a2}	$\min_{a2}$	$\max_{a2}$	$\bar{x}_{a2}$	$x_{a2;0,50}$	s_{a2}
...	...					
a_m	n_{am}	$\min_{am}$	$\max_{am}$	$\bar{x}_{am}$	$x_{am;0,50}$	s_{am}
spolu	n	min	max	$\bar{x}$	$x_{0,50}$	s

Z opisných štatistík, prípadne v širšom zábere, možno prezentovať niekoľko verzií **grafov**:

1. Graf priemerov (os x-ová obsahuje vyznačené úrovne $a_1, a_2, \dots, a_m$; os y-ová hodnoty priemerov; stĺpcový alebo spojnicový graf).
2. Graf 95-percentných intervalov spoľahlivosti strednej hodnoty pre jednotlivé úrovne faktora A .
3. Graf box-plotov (box-ploty pre jednotlivé úrovne faktora A).

4. Spojnicový graf miním, maxím, priemerov a mediánov pre jednotlivé úrovne faktora A.

Model vplyvu úrovni faktora A na variabilitu hodnôt premennej X:

$$X_{ij} = \mu + \alpha_i + e_{ij} \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, n_i)$$

kde je

X_{ij} – hodnota premennej X u j-tej jednotky na i-tej úrovni,

μ – obecná úrovňová konštanta,

$$\left(\bar{x} = \frac{1}{n} \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij} = \text{est } \mu \right)$$

α_i – efekt i-tej úrovne, $(\bar{x}_i - \bar{x}) = \text{est } \alpha_i$

$$\left(\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij} \right)$$

e_{ij} – náhodná chyba.

K overeniu **nevýznamnosti** vplyvu úrovni faktora A na variabilitu hodnôt premennej X sa testuje dvojica hypotéz:

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_m = 0$$

$$H_1 : \sum_{i=1}^m \alpha_i^2 > 0$$

V prípade homoskedasticity súboru sa k testu používa F – test, v prípade heteroskedasticity sa môže použiť neparametrický Kruskal-Wallisov test. (Autor predpokladá počítačové riešenie úlohy a preto vzorce testovacích štatistík neuvádza.)

Celková variabilita premennej X sa môže merať aj sumou štvorcov SS (SS je jeden z medzivýsledkov pri výpočte rozptylu s^2 a tým aj smerodajnej odchýlky s). **Podiel** vplyvu úrovni faktora A na celkovú variabilitu SS premennej X je R^2 (R-square) a určíme ho z rovnice:

$$SS = SS_A + SS_E, \quad \text{pričom } R^2 = SS_A/SS$$

Praktické závery formulujeme v súlade s cieľom, z dôvodu ktorého sme sa rozhodli pre analýzu rozptylu a získaných výsledkov.

Príklad – analýza vplyvu právnej formy na variabilitu výsledkov hospodárenia po zdanení v súbore firiem Výroby elektrických strojov a prístrojov

Premennú X predstavuje ukazovateľ Výsledok hospodárenia po zdanení (V64) - ZISK. Údaje za jednotlivé firmy sú v tisíc Sk. Faktorová premenná A predstavuje kód právnej formy firmy - FORMA.

K riešeniu sa použil systém Statgraphics, procedúra One-way ANOVA...
Výstup je nasledujúci:

Analysis Summary

Dependent variable: ZISK
Factor: FORMA

Number of observations: 402
Number of levels: 7

The StatAdvisor

This procedure performs a one-way analysis of variance for ZISK. It constructs various tests and graphs to compare the mean values of ZISK for the 7 different levels of FORMA. The F-test in the ANOVA table will test whether there are any significant differences amongst the means. If there are, the Multiple Range Tests will tell you which means are significantly different from which others. If you are worried about the presence of outliers, choose the Kruskal-Wallis Test which compares medians instead of means. The various plots will help you judge the practical significance of the results, as well as allow you to look for possible violations of the assumptions underlying the analysis of variance.

Z údajov vidíme, že súbor obsahuje 402 firiem a premenná FORMA 7 úrovní.

Summary Statistics for ZISK

FORMA	Count	Average	Median
101	12	266,333	103,0
102	6	731,667	321,5
111	2	-122,0	-122,0
112	344	3712,07	61,0
121	31	8969,77	1267,0
205	2	7874,5	7874,5
421	5	-1120,6	316,0
Total	402	3911,7	77,5
FORMA	Standard deviation	Minimum	Maximum
101	468,743	-125,0	1211,0
102	980,766	-6,0	2510,0
111	172,534	-244,0	0,0
112	36400,3	-337127,0	465736,0
121	33247,2	-52918,0	151962,0
205	13926,5	-1973,0	17722,0
421	11393,8	-20283,0	9277,0
Total	34940,9	-337127,0	465736,0

The StatAdvisor

This table shows various statistics for ZISK for each of the 7 levels of FORMA. The one-way analysis of variance is primarily intended to compare the means of the different levels, listed here under the Average column. Select Means Plot from the list of Graphical Options to display the means graphically.

WARNING: There is more than a 3 to 1 difference between the smallest standard deviation and the largest. This may cause problems since the analysis of variance assumes that the standard deviations at all levels are equal. Select Variance Check from the list of Tabular Options to run a formal statistical test for differences among the sigmas. You may want to consider transforming the values of ZISK to remove any dependence of the standard deviation on the mean.

Ďalšia časť výstupu obsahuje vybrané štatistiky pre jednotlivé úrovne faktora FORMA. Vidíme, že máme 12 fyzických osôb nezapísaných do obchodného registra (101), 6 fyzických osôb zapísaných do obchodného registra (102), 2 verejné obchodné spoločnosti (111), 344 spoločností s ručením obmedzeným (112), 31 akciových spoločností (121), 2 družstvá (205) a 5 zahraničných (421).

ANOVA Table for ZISK by FORMA

Analysis of Variance					
Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Between groups	1,21753E9	6	2,02921E8	0,16	0,9860
Within groups	4,88351E11	395	1,23633E9		
Total (Corr.)	4,89568E11	401			

The StatAdvisor

The ANOVA table decomposes the variance of ZISK into two components: a between-group component and a within-group component. The F-ratio, which in this case equals 0,164131, is a ratio of the between-group estimate to the within-group estimate. Since the P-value of the F-test is greater than or equal to 0,05, there is not a statistically significant difference between the mean ZISK from one level of FORMA to another at the 95,0% confidence level.

Z tabuľky analýzy rozptylu môžeme konštatovať, že faktor FORMA neovplyvňuje variabilitu hodnôt premennej ZISK (P-Value 0,9860 > 0,05). Podiel vysvetlenej variability je veľmi malý ($R^2 = 1,21753E9 / 4,89568E11$).

Na ďalšej strane je analýza kontrastov. Z výsledkov opäť môžeme konštatovať, že medzi firmami zoskupenými podľa právnej formy nie sú štatisticky významné rozdiely.

Multiple Range Tests for ZISK by FORMA

Method: 95,0 percent LSD

FORMA	Count	Mean	Homogeneous Groups
421	5	-1120,6	X
111	2	-122,0	X
101	12	266,333	X
102	6	731,667	X
112	344	3712,07	X
205	2	7874,5	X
121	31	8969,77	X

Contrast	Difference	+/- Limits
101 - 102	-465,333	34563,6
101 - 111	388,333	52796,8
101 - 112	-3445,74	20300,4
101 - 121	-8703,44	23502,4
101 - 205	-7608,17	52796,8
101 - 421	1386,93	36795,8
102 - 111	853,667	56442,1
102 - 112	-2980,41	28466,1
102 - 121	-8238,11	30831,4
102 - 205	-7142,83	56442,1
102 - 421	1852,27	41858,6
111 - 112	-3834,07	49022,2
111 - 121	-9091,77	50432,5
111 - 205	-7996,5	69127,2
111 - 421	998,6	57836,0
112 - 121	-5257,7	12963,0
112 - 205	-4162,43	49022,2
112 - 421	4832,67	31138,5
121 - 205	1095,27	50432,5
121 - 421	10090,4	33314,6
205 - 421	8995,1	57836,0

* denotes a statistically significant difference.

The StatAdvisor

This table applies a multiple comparison procedure to determine which means are significantly different from which others. The bottom half of the output shows the estimated difference between each pair of means. There are no statistically significant differences between any pair of means at the 95,0% confidence level. At the top of the page, one homogenous group is identified by a column of X's. Within each column, the levels containing X's form a group of means within which there are no statistically significant differences. The method currently being used to discriminate among the means is Fisher's least significant difference (LSD) procedure. With this method, there is a 5,0% risk of calling each pair of means significantly different when the actual difference equals 0.

Variance Check

Cochran's C test: 0,480879 P-Value = 0,0

Bartlett's test: 1,3964 P-Value = 0,0

Hartley's test: 44510,3

The StatAdvisor

The three statistics displayed in this table test the null hypothesis that the standard deviations of ZISK within each of the 7 levels of FORMA is the same. Of particular interest are the two P-values. Since the smaller of the P-values is less than 0,05, there is a statistically significant difference amongst the standard deviations at the 95,0% confidence level. This violates one of the important assumptions underlying the analysis of variance and will invalidate most of the standard statistical tests. Since the standard deviations differ by more than a factor 3 to 1 and the sample sizes are not equal, the P-values and significance levels of the tests may be off significantly.

Z kontroly homoskedasticity môžeme podľa oboch P-hodnôt konštatovať rozdielne hodnoty rozptylov pre jednotlivé podsúbory firiem špecifikované rovnakou právnou formou. Z tohto záveru vyplýva potreba použiť neparametrický test.

Kruskal-Wallis Test for ZISK by FORMA

FORMA	Sample Size	Average Rank
101	12	192,708
102	6	240,167
111	2	95,25
112	344	197,949
121	31	241,887
205	2	204,5
421	5	211,4

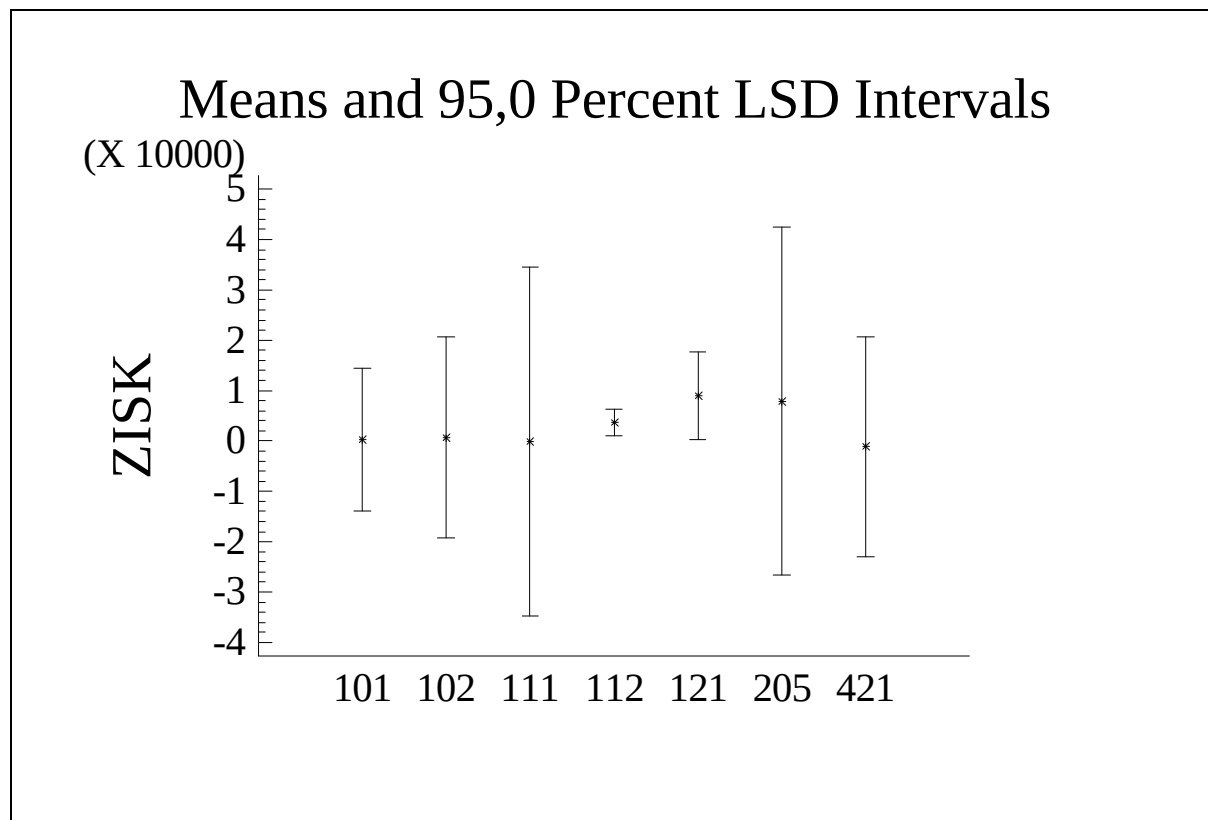
Test statistic = 6,51051 P-Value = 0,368492

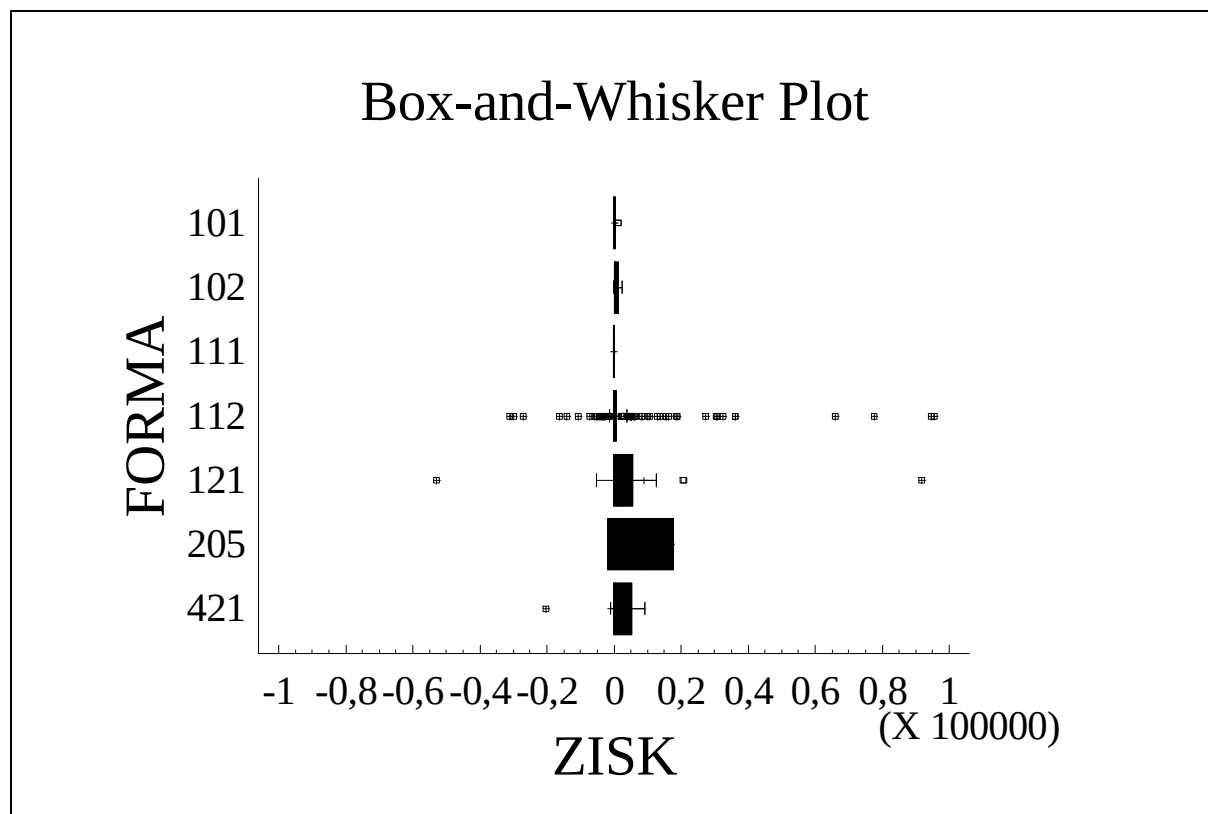
The StatAdvisor

The Kruskal-Wallis test tests the null hypothesis that the medians of ZISK within each of the 7 levels of FORMA are the same. The data from all the levels is first combined and ranked from smallest to largest. The average rank is then computed for the data at each level. Since the P-value is greater than or equal to 0,05, there is not a statistically significant difference amongst the medians at the 95,0% confidence level.

Aj podľa Kruskal-Wallisovko testu môžeme konštatovať štatisticky nevýznamný vplyv úrovni faktoru FORMA na variabilitu hodnôt premennej ZISK.

Aj grafická prezentácia 95 percentných intervalov spoľahlivosti pre jednotlivé úrovne faktora FORMA naznačuje nevýznamný vplyv úrovni na variabilitu zisku.





Záver

Zisky firiem sa pohybujú okolo nuly. Menší počet jednotiek v súbore znamená väčšiu variabilitu odhadu strednej hodnoty. Pri analýze zisku nemusíme brať do úvahy právnu formu organizácie.

Literatúra

Chajdiak, J.: Štatistické úlohy a ich riešenie v Exceli. Bratislava, Statis 2005

Chajdiak, J. – Komorník, J. – Komorníková, M.: Štatistické metódy. Bratislava, Statis 1999

Stankovičová I., Vojtková M.: Viacrozmerné štatistické metódy pre ekonómov a manažérov. IURA, Bratislava 2007.

Kanderová, M. – Úradníček, V.: Štatistika a pravdepodobnosť pre ekonómov. 1. časť. OZ Financ, Banská Bystrica 2005.

Kanderová, M. – Úradníček, V.: Štatistika a pravdepodobnosť pre ekonómov. 2. časť. OZ Financ, Banská Bystrica 2007.

Kontakt na autora:

chajdiak@statis.biz

Príspevok bol spracovaný v rámci riešenia grantovej úlohy VEGA 1/2631/05 „Analýza možností aplikácie viacrozmerných štatistických metód na skúmanie ekonomických výsledkov na príklade priemyslu SR prípadne iných oblastí ekonomiky“.

Vývoj vybraných ukazovateľov priemyslu podľa intervalov vytvorenej pridanej hodnoty v rokoch 1995, 2000 a 2006

Doc. Ing. Jozef Chajdiak, CSc., STATIS
Ing. Edita Holičková, Štatistický úrad SR

Abstract

Contribution contains evaluation of development chosen indicators in industry SR in years 1995, 2000 and 2006.

V priemysle Slovenskej republiky v roku 2006 pôsobilo 2367 podnikateľských subjektov s počtom nad 20 zamestnancov. Ich ekonomické výsledky, použité z predbežných štvrťročných zisťovaní sú porovnávané s výsledkami, dosiahnutými v rokoch 1995 a 2000.

Zatiaľ čo obdobie rokov 1995-2000 bolo poznačené postupom demonopolizácie a atomizácie priemyslu, v nasledujúcich rokoch sa počet podnikateľských subjektov vcelku stabilizoval, zmiernilo sa tempo znižovania zamestnanosti pri raste objemu mzdových nákladov. Výnosy sa v roku 2006 oproti roku 2000 zvýšili o 103,6 % a hospodársky výsledok (pred zdanením) významne vzrástol.

Súhrnné výsledky priemyslu SR

Tab. 1 v mil. Sk v bežných cenách

Ukazovateľ/Rok	2006	2004	2000	1995	Index 2006/2000 0	Index 2000/1995
Počet podnikov	2 367	2 127	2 144	1895	99,2	113,1
Počet zamestnancov - osoby	428 309	413 168	437 912	562 570	94,3	77,8
Pridaná hodnota	352 426	309 462	209 786	148 154	147,5	141,6
Výnosy	1 932 523	1 420 360	949 068	581 117	149,6	163,3
Hospodársky výsledok	140 580	96 289	14 462	33 017	665,8	43,8

Pri znižovaní zamestnanosti v priemysle vzrástla produktivita práce, meraná vytvorenou pridanou hodnotou v priemere na jedného zamestnanca v roku 2000 v porovnaní s rokom 1995 o 82,1 % na 479 tis. Sk, v roku 2006 sa produktivita práce v porovnaní s rokom 2000 zvýšila o 71,8 % na 823 tis. Sk (v bežných cenách). Celkové výnosy na jedného zamestnanca v priemysle sa v období rokov 1995-2000 zdvojnásobili, v roku 2006 sa zvýšili v porovnaní s rokom 2000 o 108,2 %.

Zatiaľ čo vývoj vybraných ukazovateľov efektívnosti priemyslu v prvej etape hodnotenia (1995-2000) ovplyvnila nižšia intenzita tvorby hospodárskeho výsledku, súvisiaca s problematikou reštrukturalizácie, ako aj nižšie tempo investovania do výroby s vyšším zhodnocovaním vstupov, v roku 2006 v porovnaní s rokom 2000 sa pri miernejšom tempe rastu zvýšila tvorba pridanej hodnoty a tvorba výnosov na zamestnanca. Výraznejšie vzrástla rentabilita výnosov, meraná podielom (nezdaneného) hospodárskeho výsledku na výnosoch.

Vybrané ukazovatele ekonomickej efektívnosti v priemysle

Tab. 2

Ukazovateľ/Rok	2006	2000	1995	Index 2006/2000	Index 2000/1995
Pridaná hodnota v tis. Sk / zamestnanca	823	479	263	171,8	182,1
Výnosy v tis. Sk / zamestnanca	4 512	2 167	1 033	208,2	209,8
Podiel pridanej hodnoty v hal. / 1 Sk výnosov ¹⁾	18,20	22,10	25,50	-3,90	- 3,40
Podiel hospodárskeho výsledku v hal. /1 Sk výnosov ¹⁾	7,30	1,50	5,70	5,80	-4,20

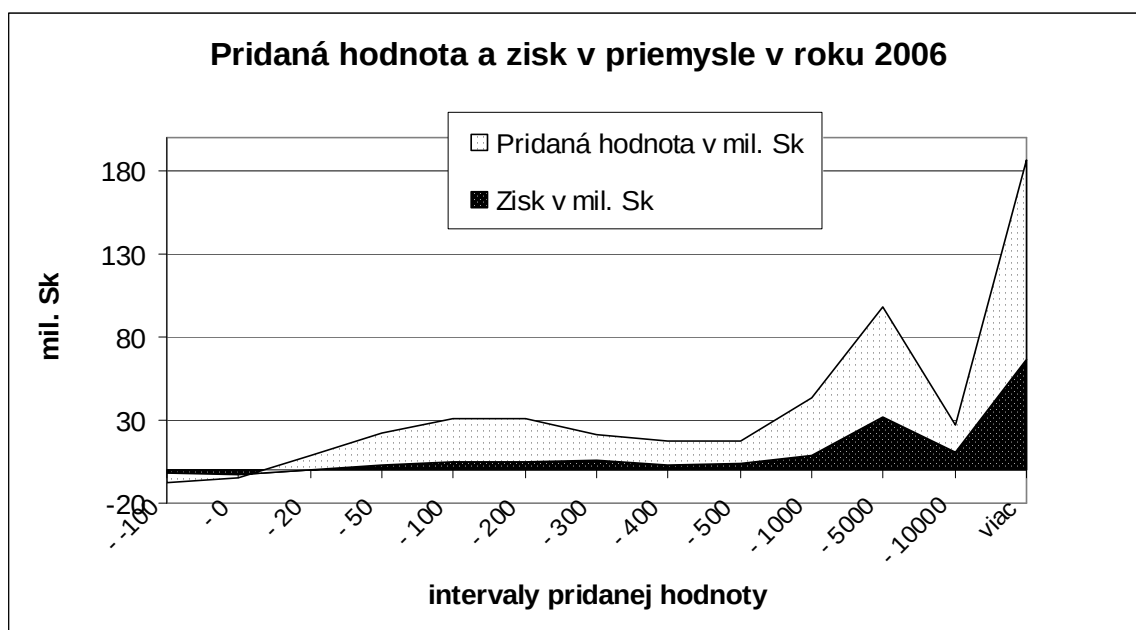
¹⁾ v stĺpci „Index“ sú uvádzané rozdiely príslušných hodnôt

Na 1 Sk výnosov v roku 2006 pripadlo

- 18,20 halierov pridanej hodnoty, pričom tento podiel v porovnaní s rokom 2000 poklesol o 3,90 haliera
- 7,30 haliera zisku (v porovnaní s rokom 2000 tento podiel vzrástol o 5,80 haliera)

V roku 2006 v priemere na jeden priemyselný podnik pripadalo 181 zamestnancov (v porovnaní s rokom 2000 o 23 a v porovnaní s rokom 1995 o 115 zamestnancov menej). Z hľadiska tvorbu pridanej hodnoty v roku 2006 pripadlo na jeden podnik v priemere 148,9 mil. Sk pridanej hodnoty (97,8 mil. Sk v roku 2000 a 78,2 mil. Sk v roku 1995).

Tento priemer skreslili výsledky 105 podnikov, v ktorých bola negatívna, resp. nulová tvorba pridanej hodnoty. Táto skupina „neefektívnych“ podnikov v priemysle v roku 2006 participovala 1,5 % na zamestnanosti odvetvia, hospodársky výsledok odvetvia znížila o 1,7 % a tvorbu pridanej hodnoty o 2,2 %, na tvorbe výnosov sa podieľala 0,4 %. Podiel dopad tejto skupiny podnikov na tvorbu hospodárskeho výsledku priemyslu bol v porovnaní s rokom 2000 menej výrazný (v roku 2000 hospodársky výsledok priemyslu znížila o 21 %).



Početne najväčšia skupina 899 podnikov s pomerne nízkou priemernou pridanou hodnotou 9,8 mil. Sk (v intervale od 1 do 20 mil. Sk) roku 2006 vytvorila 2,5 % pridanej hodnoty priemyslu, 2,7 % výnosov, 0,6 % hospodárskeho výsledku a zamestnávala 9,8 % zamestnancov.

Ďalšia početne významná skupina 580 podnikov s priemernou tvorbou pridanej hodnoty 32,1 mil. Sk (v intervale od 21 do 50 mil. Sk) v roku 2006 vytvorila 5,3 % pridanej hodnoty

v priemysle, 4,3 % výnosov, podieľala sa 11,7 % na zamestnanosti a 2,3 % na tvorbe hospodárskeho výsledku.

V intervale od 51 do 100 mil. Sk bolo v roku 2006 podľa vytvorenej pridanej hodnoty 359 podnikov, vytvorili 7,4 % pridanej hodnoty v priemysle, 6,7 % výnosov a 3,7 % zisku. Na zamestnanosti v priemysle participovali 13,4 %.

V skupine podnikov s vytvorenou pridanou hodnotou v intervale od 101 do 200 mil. Sk v roku 2006 bolo 192 priemyselných podnikov. Na tvorbe pridanej hodnoty v priemysle sa podieľali 7,5 %, na tvorbe výnosov 6,7 %, na tvorbe zisku 3,4 % a na zamestnanosti v priemysle participovali 13,4 %.

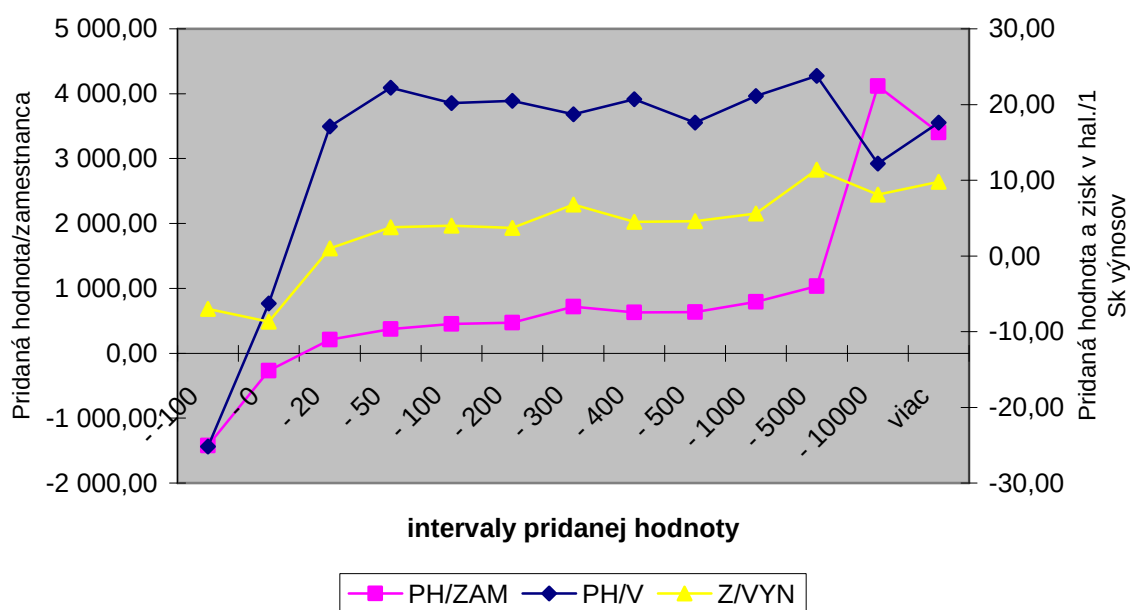
Porovnanie vybraných ukazovateľov efektívnosti v skupinách podnikov s najvyššou a najnižšou efektívnosťou podľa intervalov pridanej hodnoty

Tab. 3

Ukazovateľ	Rok 2006			Rok 2000		
	Spolu	0 a menej	1000 mil. Sk a viac	Spolu	0 a menej	1000 mil. Sk a viac
Pridaná hodnota v tis. Sk / zamestn.	823	-1 696	8 556	479	-103	1 159
Pridaná hodnota v hal./1 Sk výnosov	18,20	-6,30	23,80	22,10	-15,60	22,20
Hospodársky vý- sledok v hal. / 1 Sk výnosov	7,30	-8,70	11,40	1,50	-51,10	3,60

Všetky priemyselné podniky s tvorbou pridanej hodnoty v intervale do 200 mil. Sk participujú na celkovom počte priemyselných subjektov 90,3 %, na tvorbe pridanej hodnoty 20,5 %, na tvorbe výnosov 23 %, zamestnávali polovicu (50,4 %) zamestnancov v priemysle a vytvorili 6,8 % zisku priemyslu.

Vybrané ukazovatele efektívnosti priemyslu v roku 2006



Najpozitívnejšie ovplyvňovala vývoj priemyslu skupina 95 podnikov (v intervale tvorby pridanej hodnoty nad 1 mil. Sk) s 57,3 % podielom na pridanej hodnote priemyslu, 56,6 % podielom na tvorbe výnosov a 24 % podielom na zamestnanosti. Táto najvýznamnejšia skupina podnikov vytvorila 77,8 % zisku celého priemyslu. Je v nej zaradených 8 firiem s najvyššou tvorbou pridanej hodnoty, participujúcich 34 % na objeme pridanej hodnoty odvetvia, 35,2 % na výnosoch, 47,5 % na zisku a 9,1 % na zamestnanosti v priemysle.

Z porovnania vývoja vybraných ukazovateľov efektívnosti v roku 2006 a 2000 vyplýva pretrvávajúci negatívny vplyv podnikov s tvorbou pridanej hodnoty v pásme nulovom až mínusovom na efektívnosť priemyslu. Aj keď sa počet týchto subjektov v porovnávanom období znížil, ich negatívny dopad na efektívnosť priemyslu je zrejмый. Vplyv podnikov s najvyššou tvorbou pridanej hodnoty na ekonomické výsledky priemyslu sa v porovnávanom období zvýraznil.

Tabuľkové prílohy

Rozdelenie podnikov podľa pridanej hodnoty v roku 2006

Trieda	Priemer		N	fN	kN	kfN	pPH	PH	fPH	kPH	kfPH	pVYN	VYN	fVYN	kVYN	kfVYN	pZISK	ZISK	fZISK	kZISK	kfZISK
- -100	-369,6		16	0,7	16	0,7	-369,6	-5914	-1,7	-5914	-1,7	1466	23451	1,2	23451	1,2	-102,5	-1640	-1,2	-1640	-1,2
- 0	-19,1		91	3,8	107	4,5	-19,1	-1741	-0,5	-7655	-2,2	305	27771	1,4	51222	2,7	-26,5	-2410	-1,7	-4050	-2,9
- 20	9,8		899	38,0	1006	42,5	9,8	8785	2,5	1130	0,3	57	51339	2,7	102561	5,3	0,6	497	0,4	-3553	-2,5
- 50	32,1		580	24,5	1586	67,0	32,1	18616	5,3	19746	5,6	145	83819	4,3	186380	9,6	5,5	3211	2,3	-342	-0,2
- 100	72,7		359	15,2	1945	82,2	72,7	26112	7,4	45858	13,0	360	129372	6,7	315752	16,3	14,4	5180	3,7	4838	3,4
- 200	137,3		192	8,1	2137	90,3	137,3	26366	7,5	72224	20,5	671	128927	6,7	444679	23,0	24,8	4762	3,4	9600	6,8
- 300	245,9		63	2,7	2200	92,9	245,9	15493	4,4	87717	24,9	1314	82800	4,3	527479	27,3	88,8	5593	4,0	15193	10,8
- 400	344,9		41	1,7	2241	94,7	344,9	14139	4,0	101856	28,9	1667	68338	3,5	595817	30,8	75,8	3106	2,2	18299	13,0
- 500	452,3		31	1,3	2272	96,0	452,3	14022	4,0	115878	32,9	2573	79762	4,1	675579	35,0	118,3	3668	2,6	21967	15,6
- 1000	662,8		52	2,2	2324	98,2	662,8	34464	9,8	150342	42,7	3141	163319	8,5	838898	43,4	176,9	9199	6,5	31166	22,2
- 5000	1888,0		35	1,5	2359	99,7	1888,0	66079	18,7	216421	61,4	7919	277163	14,3	1116061	57,8	906,6	31732	22,6	62898	44,7
- 10000	5498,0		3	0,1	2362	99,8	5498,0	16494	4,7	232915	66,1	45187	135560	7,0	1251621	64,8	3643,7	10931	7,8	73829	52,5
viac	23902,2		5	0,2	2367	100,0	23902,2	119511	33,9	352426	100,0	136181	680903	35,2	1932524	100,0	13350,2	66751	47,5	140580	100,0
Spolu	148,9		2 367	100,0			148,9	352 426	100,0			816,4	1 932 524	100,0			59,39	140 580	100,0		

Trieda	Priemer	pOBRAT	OBRAT	fOBRAT	kOBRAT	kfOBRAT	pCK	CK	fCK	kCK	kfCK	pVI	VI	fVI	kVI	kfVI	pZAM	ZAM	fZAM	kZAM	kfZAM
- -100	-369,6	933	14924	0,9	14924	0,9	2973,4	47575	2,9	47575	2,9	1052,9	16847	1,8	16847	1,8	259,4	4150	1,0	4150	1,0
- 0	-19,1	266	24209	1,5	39133	2,4	214,9	19554	1,2	67129	4,1	45,9	4181	0,5	21028	2,3	70,6	6428	1,5	10578	2,5
- 20	9,8	46	41694	2,5	80827	4,9	37,5	33739	2,0	100868	6,1	11,8	10635	1,2	31663	3,5	46,5	41801	9,8	52379	12,2
- 50	32,1	126	72860	4,4	153687	9,3	107,3	62236	3,8	163104	9,9	39,5	22927	2,5	54590	6,0	86,4	50128	11,7	102507	23,9
- 100	72,7	316	113442	6,9	267129	16,1	246,7	88568	5,4	251672	15,3	97,1	34869	3,8	89459	9,8	160,3	57554	13,4	160061	37,4
- 200	137,3	593	113799	6,9	380928	23,0	453,9	87140	5,3	338812	20,6	195,3	37491	4,1	126950	13,9	289,8	55649	13,0	215710	50,4
- 300	245,9	1092	68799	4,2	449727	27,2	885,3	55775	3,4	394587	23,9	465,0	29297	3,2	156247	17,1	343,0	21607	5,0	237317	55,4
- 400	344,9	1426	58483	3,5	508210	30,7	1175,4	48190	2,9	442777	26,9	454,7	18644	2,0	174891	19,1	548,4	22486	5,2	259803	60,7
- 500	452,3	2338	72468	4,4	580678	35,1	1581,3	49019	3,0	491796	29,8	655,5	20321	2,2	195212	21,3	710,7	22032	5,1	281835	65,8
- 1000	662,8	2789	145035	8,8	725713	43,8	2230,4	115981	7,0	607777	36,9	1231,7	64047	7,0	259259	28,3	836,2	43484	10,2	325319	76,0
- 5000	1888,0	6718	235129	14,2	960842	58,0	12294,3	430300	26,1	1038077	63,0	7656,3	267970	29,3	527229	57,6	1825,1	63877	14,9	389196	90,9
- 10000	5498,0	42137	126410	7,6	1087252	65,7	15227,7	45683	2,8	1083760	65,8	7544,7	22634	2,5	549863	60,1	1335,3	4006	0,9	393202	91,8
viac	23902,2	113767	568833	34,3	1656085	100,0	112845,6	564228	34,2	1647988	100,0	73064,2	365321	39,9	915184	100,0	7021,6	35108	8,2	428310	100,0
Spolu	148,89	699,7	1 656 085	100			696,2	1 647 988	100,0			386,6	915 184	100,0			181,0	428 310	100,0		

N	počet firiem	ZISK	výsledok hospodárenia(v mil. Sk)	f	relatívna početnosť ukazovateľa (v percentách)
VI	vlastné imanie	CK	celkový kapitál (pasíva spolu) (v mil. Sk)	k	kumulatívne početnosť ukazovateľa
PH	pridaná hodnota (v mil. Sk)	VS	výrobná spotreba (v mil. Sk)	p	priemerná hodnota ukazovateľa
VYN, V	výnosy (v mil. Sk)	SPMAT	spotreba materiálu (v mil. Sk)		
OBRAT	tržby za vlastné výkony a tovar (v mil. Sk)	SLUŽBY	služby (v mil. Sk)		
ZAM	zamestnanci (priem.evid.počet)	ODP	odpisy (v mil. Sk)		

Frekvenčná tabuľka: Pridaná hodnota (SR Priemysel Prod 3-04, rok 2006, v mil. Sk)

Príloha 1/2

Trieda	Priemer	pVS	VS	fVS	kVS	kfVS	pSPMAT	SPMAT	fSPMAT	kSPMAT	kfSPMAT	pSLUZBY	SLUZBY	fSLUZBY	kSLUZBY	kfSLUZBY	pODP	ODP	fODP	kODP	kfODP
- -100	-369,6	1183,3	18932	1,5	18932	1,5	923,8	14780	1,4	14780	1,4	259,56	4153	2,4	4153	2,4	54,8	877	1,2	877	1,2
- 0	-19,1	259,8	23642	1,9	42574	3,4	212,5	19340	1,8	34120	3,2	47,29	4303	2,5	8456	4,8	8,7	789	1,1	1666	2,3
- 20	9,8	33,6	30176	2,4	72750	5,9	26,6	23929	2,2	58049	5,5	6,95	6247	3,6	14703	8,4	1,8	1592	2,2	3258	4,6
- 50	32,1	86,7	50291	4,1	123041	9,9	66,2	38390	3,6	96439	9,1	20,52	11901	6,8	26604	15,2	5,1	2940	4,1	6198	8,7
- 100	72,7	222,8	79978	6,5	203019	16,4	181,2	65034	6,1	161473	15,2	41,63	14944	8,6	41548	23,8	12,4	4442	6,2	10640	15,0
- 200	137,3	395,7	75974	6,1	278993	22,5	324,6	62331	5,9	223804	21,0	71,06	13643	7,8	55191	31,6	23,3	4477	6,3	15117	21,3
- 300	245,9	820,9	51716	4,2	330709	26,7	708,0	44607	4,2	268411	25,2	112,83	7108	4,1	62299	35,7	50,8	3203	4,5	18320	25,8
- 400	344,9	1058,3	43392	3,5	374101	30,2	841,5	34500	3,2	302911	28,4	216,88	8892	5,1	71191	40,8	68,6	2813	4,0	21133	29,7
- 500	452,3	1827,7	56659	4,6	430760	34,8	1533,0	47524	4,5	350435	32,9	294,68	9135	5,2	80326	46,0	96,5	2991	4,2	24124	33,9
- 1000	662,8	2038,4	105997	8,6	536757	43,3	1728,3	89872	8,4	440307	41,4	310,10	16125	9,2	96451	55,3	126,2	6564	9,2	30688	43,2
- 5000	1888,0	4706,2	164717	13,3	701474	56,6	3564,9	124773	11,7	565080	53,1	1141,23	39943	22,9	136394	78,2	403,5	14121	19,9	44809	63,0
- 10000	5498,0	36641,7	109925	8,9	811399	65,5	33325,0	99975	9,4	665055	62,5	3317,00	9951	5,7	146345	83,9	772,0	2316	3,3	47125	66,3
viac	23902,2	85574,8	427874	34,5	1239273	100,0	79941,0	399705	37,5	1064760	100,0	5633,80	28169	16,1	174514	100,0	4796,2	23981	33,7	71106	100,0
Spolu	148,9	523,6	1 239 273	100,0			449,8	1 064 760	100,0			73,7	174 514	100,0			30,0	71 106	100,0		

Frekvenčná tabuľka: Pridaná hodnota (SR Priemysel Prod 3-04, rok 2006, v mil. Sk)

Trieda	Priemer	Z/VI	PH/ZAM	V/CK	Z/V	PH/V	VS/V	SPMAT/V	SLUZBY/V	ODP/V	CK/VI	CK/ZAM
- -100	-369,6	-9,7	-1425,06	0,49	-7,0	-25,2	80,73	63,03	17,7	3,7	2,82	11 464
- 0	-19,1	-57,6	-270,85	1,42	-8,7	-6,3	85,13	69,64	15,5	2,8	4,68	3 042
- 20	9,8	4,7	210,16	1,52	1,0	17,1	58,78	46,61	12,2	3,1	3,17	807
- 50	32,1	14,0	371,37	1,35	3,8	22,2	60,00	45,80	14,2	3,5	2,71	1 242
- 100	72,7	14,9	453,70	1,46	4,0	20,2	61,82	50,27	11,6	3,4	2,54	1 539
- 200	137,3	12,7	473,79	1,48	3,7	20,5	58,93	48,35	10,6	3,5	2,32	1 566
- 300	245,9	19,1	717,04	1,48	6,8	18,7	62,46	53,87	8,6	3,9	1,90	2 581
- 400	344,9	16,7	628,79	1,42	4,5	20,7	63,50	50,48	13,0	4,1	2,58	2 143
- 500	452,3	18,1	636,44	1,63	4,6	17,6	71,04	59,58	11,5	3,7	2,41	2 225
- 1000	662,8	14,4	792,57	1,41	5,6	21,1	64,90	55,03	9,9	4,0	1,81	2 667
- 5000	1888,0	11,8	1034,47	0,64	11,4	23,8	59,43	45,02	14,4	5,1	1,61	6 736
- 10000	5498,0	48,3	4117,32	2,97	8,1	12,2	81,09	73,75	7,3	1,7	2,02	11 404
viac	23902,2	18,3	3404,10	1,21	9,8	17,6	62,84	58,70	4,1	3,5	1,54	16 071
Spolu	148,89	15,4	822,83	1,17	7,3	18,2	64,13	55,10	9,0	3,7	1,80	3 848

Z/VI rentabilita vlastného imania (v hal. Zisku na 1 Sk vlastného imania)
PH/ZAM produktivita práce (v 1000 Sk PH na 1 zamestnanca)
V/CK produktivita celkového kapitálu (v Sk výnosov na 1 Sk CK)
Z/V ziskovosť (v hal. zisku na 1 Sk výnosov)
PH/V podiel pridanej hodnoty na výnosoch (v hal. PH na 1 Sk výnosov)

VS/V podiel výrobnnej spotreby na výnosoch (v hal. VS na 1 Sk výnosov)
SPMAT/V podiel spotreby materiálu na výnosoch (v hal. NU na 1 Sk výnosov)
SLUZBY/V podiel služieb na výnosoch (v hal. ODP na 1 Sk výnosov)
ODP/V podiel odpisov na výnosoch (v hal. ODP na 1 Sk výnosov)
CK/VI vybavenosť vlastného imania celkovým kapitálom (v Sk CK na 1 Sk VI)
CK/ZAM vybavenosť zamestnancov celkovým kapitálom (v 1000 Sk CK na 1 zamestnanca)

VYPOVEDACIA SCHOPNOSŤ FIREMNÉHO INDEXU TRANSPARENTNOSTI

Jozef Chajdiak, Miroslav Vester

Abstract

Infin's Transparency Index (ITI) is trying to measure company's official publication activities in publicly available information sources. Company's performance measures such as profit to sales ratio can show medium strong statistical dependency to ITI. In this article preliminary results of complex research are presented.

Spoločnosť INFIN s.r.o. disponuje rozsiahlymi súbormi rôznorodých údajov. Okrem číselných údajov (hlavne o ekonomických výsledkoch firiem) sú to aj údaje najrozličnejšieho charakteru. Bola vyslovená pracovná hypotéza, že ak firma zverejňuje o sebe viac údajov, faktov a správ, jej ekonomické výsledky firmy sú lepšie.

Na posúdenie intenzity miery zverejňovania informácií o firme sme zostrojili index transparentnosti. Jeho verzia X1 pozostáva zo súčtu zverejnených informácií z oblastí:

- Výzvy na upisovanie nových akcií,
- Správy o hospodárskych výsledkoch,
- Emisné podmienky,
- Opravy,
- Údaje z registra nadácií,
- Štátna pomoc,
- Správy o hospodárení,
- Údaje z registra registrových úradov,
- Rozhodnutia o privatizácii,
- Oznámenia,
- Prospekty cenných papierov,
- Rozhodnutia o zápisoch do obchodného registra,
- Verejná súťaž,
- Oznámenia štátnych orgánov,
- Valné zhromaždenia,
- Dražby a predaj majetku,
- Účtovné závierky,
- Bez kategórie,
- Finančné dáta 1 a
- Finančné dáta 2.

V každej z uvedených oblastí môže, ale aj nemusí, byť zverejnená informácia o firme. Počet všetkých zverejnených informácií vo všetkých uvedených oblastiach za každú firmu predstavuje hodnotu premennej X1.

V roku 2005 v databázach firmy INFIN s.r.o. boli uložené informácie, okrem iných, o 3684 identifikovaných firmách, o ktorých sú známe aj číselné ekonomické výsledky za rok 2005. Spolu všetkých takýchto informácií bolo 16750, t.j. 4,5 informácie na jednu takúto firmu. Rozdelenie hodnôt premennej X1 podľa počtu informácií o firme je uvedené v tab. 1. Vidíme, že bolo 155 firiem, o ktorých je len číselná informácia o ekonomických výsledkoch (X1=1). Najčastejšie (847 krát) sa vyskytuje hodnota X1=3, t.j. jedna číselná informácia

o ekonomických výsledkoch plus dve ďalšie informácie z vyššie uvedených oblastí. Najviac o jednej firme bolo zachytených 27 informácií ($X1=27$).

Preukaznosť vypovedacej schopnosti indexu transparentnosti $X1$ sme overovali pomocou vzájomnej závislosti medzi hodnotami $X1$ a vybranými ekonomickými ukazovateľmi.

Z ekonomických ukazovateľov sme vybrali:

absolútne ukazovatele:

- výsledok hospodárenia po zdanení (r.64),
- pridaná hodnota (r.11),
- obrat ($r.01+r.05$),
- výnosy spolu,

relatívne ukazovatele:

- ziskovosť ($100 \cdot r.64 / \text{výnosy spolu}$),
- finančná produktivita práce ($r.11 / r.12$),
- rentabilita vlastného imania ($100 \cdot r.64 / r.066$),
- účinnosť celkového kapitálu ($\text{výnosy spolu} / r.065$),
- ziskovosť obratu ($100 \cdot r.29 / \text{obrat}$),
- úrokové krytie ($(\text{zisk pred zdanením} + r.41) / r.41$).

Za jednotlivé absolútne ukazovatele bola vypočítaná priemerná hodnota príslušného ukazovateľa na jednu firmu resp. medián v triede firiem určenej konkrétnou hodnotou indexu transparentnosti $X1$. Zistené výsledky sú prezentované na obr. 1 až obr.4. Cez zobrazené priemerné hodnoty príslušného ukazovateľa pre jednotlivé hodnoty indexu transparentnosti $X1$ (hodnoty väčšie alebo rovné 15 boli transformované na hodnotu 15+), sme preložili regresnú krivku (priamku resp. exponenciálnu funkciu) a vypočítali hodnotu R^2 (sú uvedené v hornej časti obrázku).

Jednotlivé relatívne ukazovatele sme vypočítali ako podiel úhrnov hodnôt čitateľa a menovateľa príslušného ukazovateľa pre triedy firiem určené konkrétnou hodnotou indexu transparentnosti $X1$. Zistené výsledky sú prezentované na obr. 5 až obr.10. Cez zobrazené priemerné hodnoty príslušného ukazovateľa pre jednotlivé hodnoty indexu transparentnosti $X1$ sme opäť preložili regresnú krivku a vypočítali hodnotu R^2 (sú uvedené v hornej časti obrázku).

Z údajov na jednotlivých obrázkoch môžeme konštatovať, že vyššia miera transparentnosti o firme (vyššia hodnota indexu transparentnosti $X1$) znamená lepšie ekonomické výsledky príslušnej firmy pri absolútnych ukazovateľoch až vo vysokom stupni štatistickej závislosti a u relatívnych ukazovateľoch v miernom až v strednom stupni štatistickej závislosti.

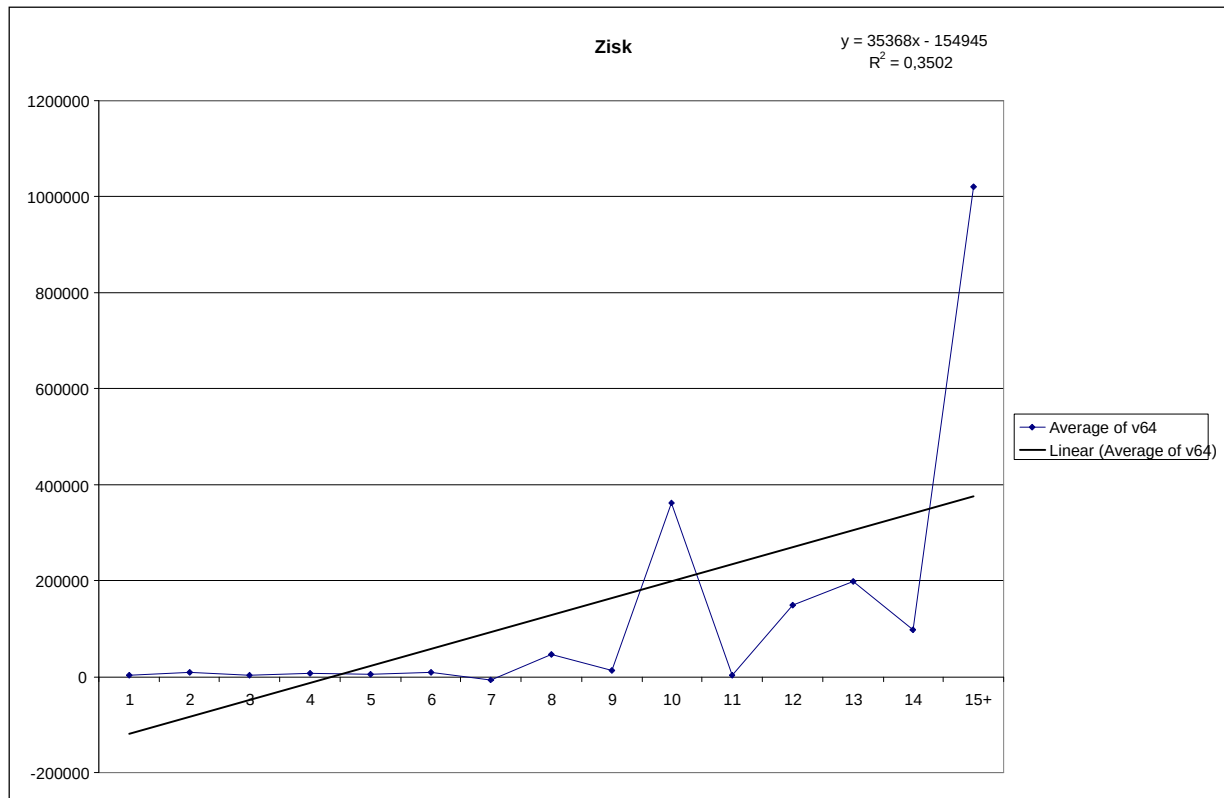
Adresa autorov:

chajdiak@statis.biz

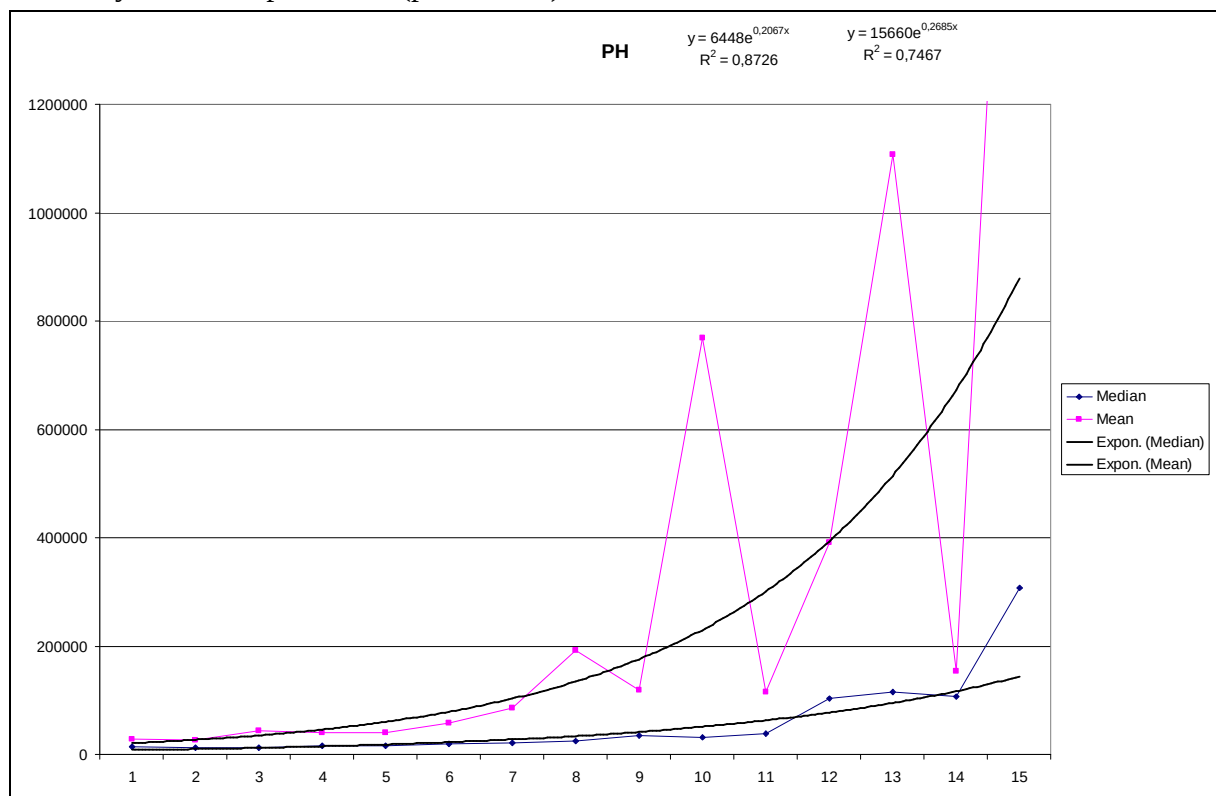
vester@infin.sk

Tab.1 Rozdelenie počtu firiem podľa dosiahnutej hodnoty X1

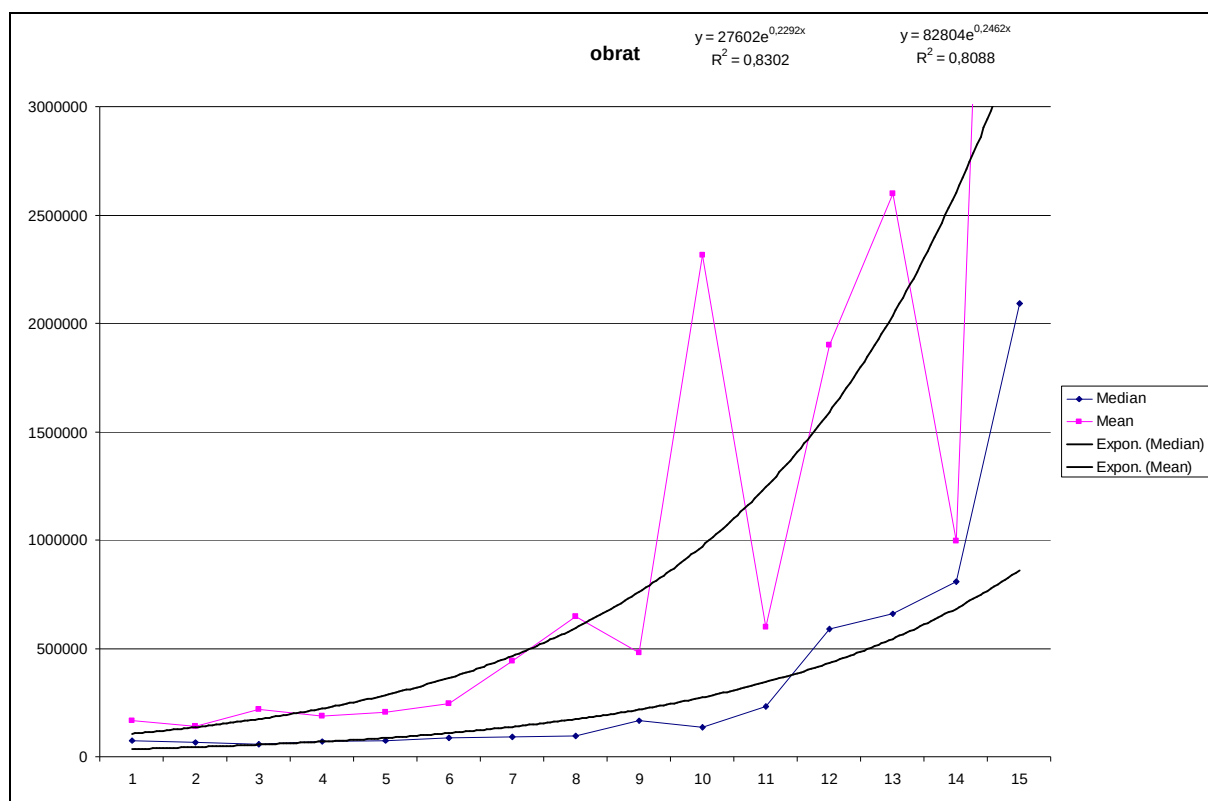
X1	Počet firiem	Počet záznamov
1	155	155
2	516	1032
3	847	2541
4	657	2628
5	555	2775
6	327	1962
7	214	1498
8	148	1184
9	87	783
10	61	610
11	28	308
12	24	288
13	22	286
14	11	154
15	11	165
16	7	112
17	5	85
18	3	54
19	3	57
22	1	22
24	1	24
27	1	27
Spolu	3684	16750



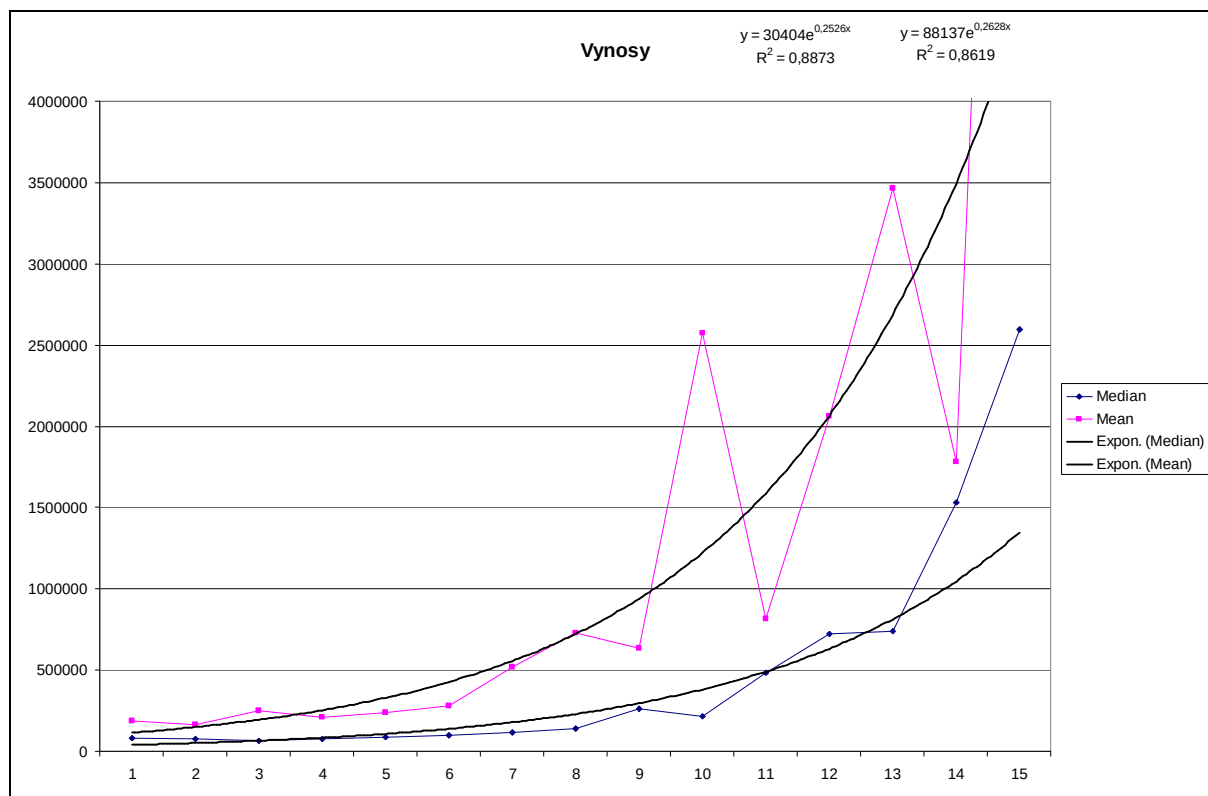
Obr.1 Výsledok hospodárenia (po zdanení)



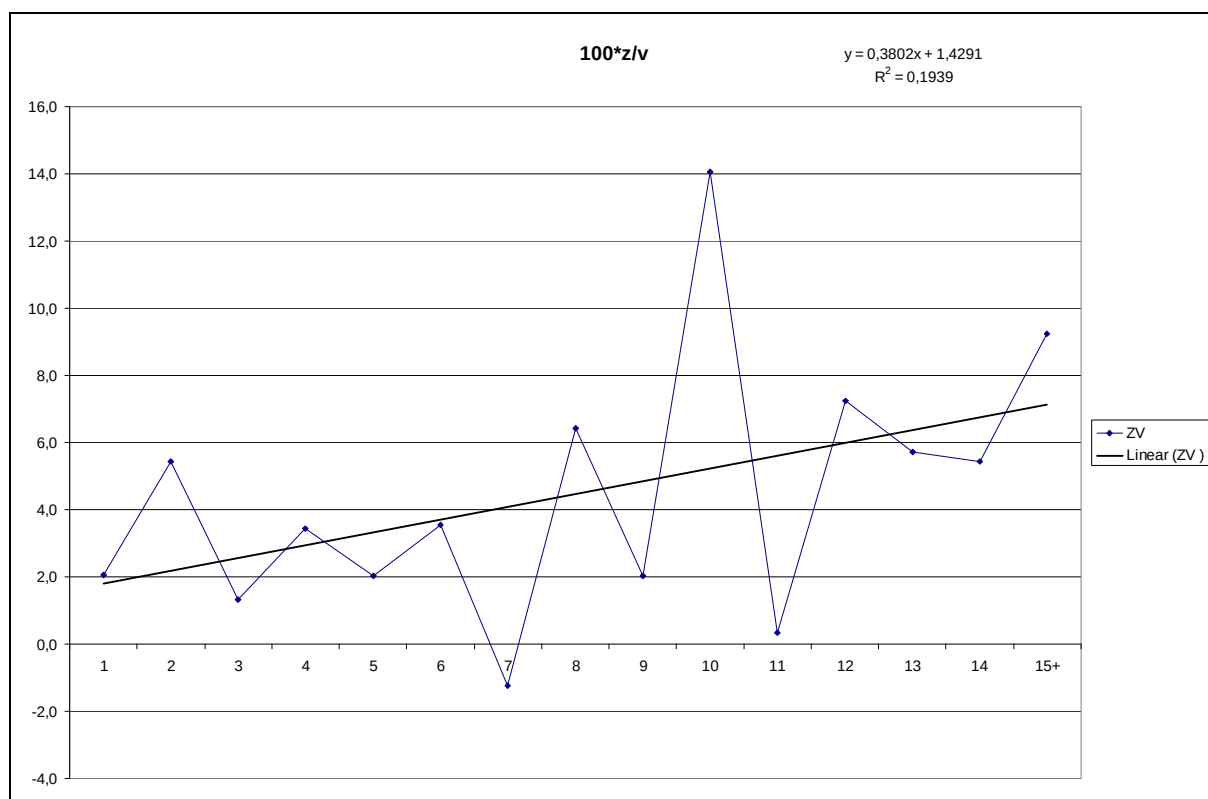
Obr.2 Pridaná hodnota



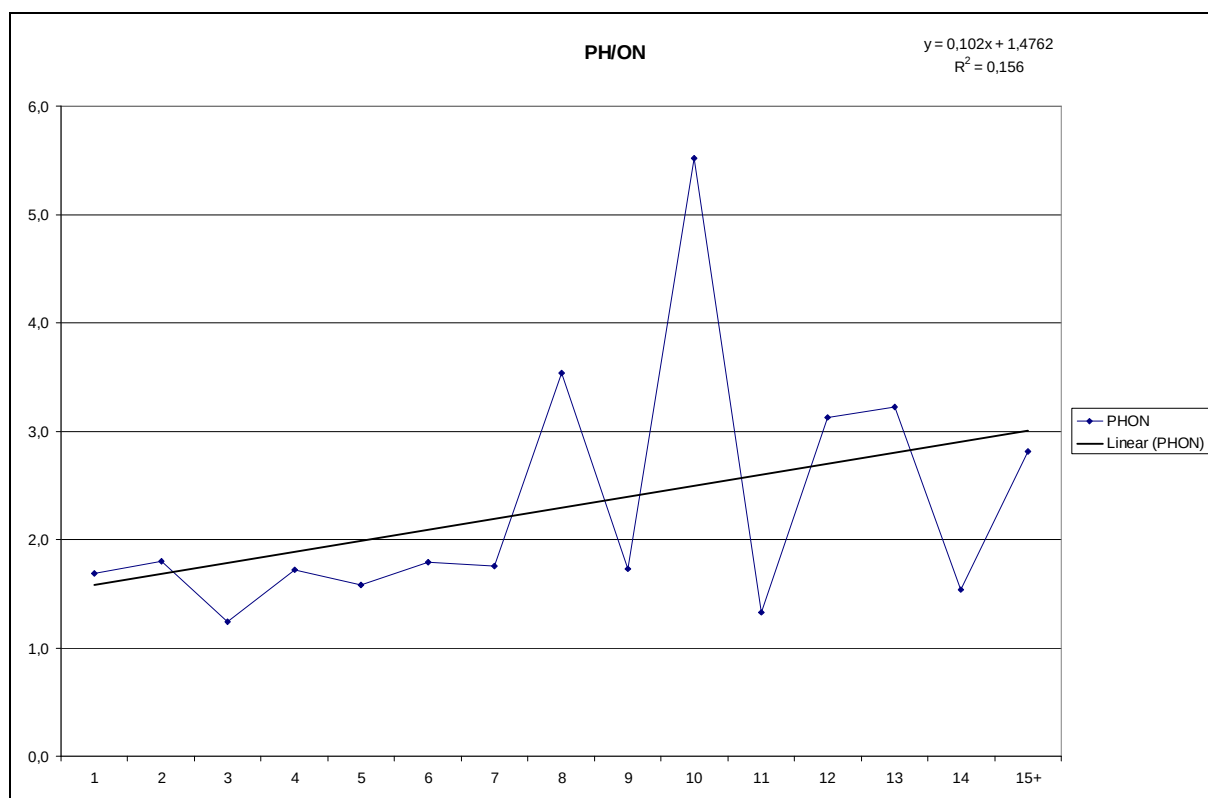
Obr. 3 Obrat (tržby za vlastní výkony a tovar)



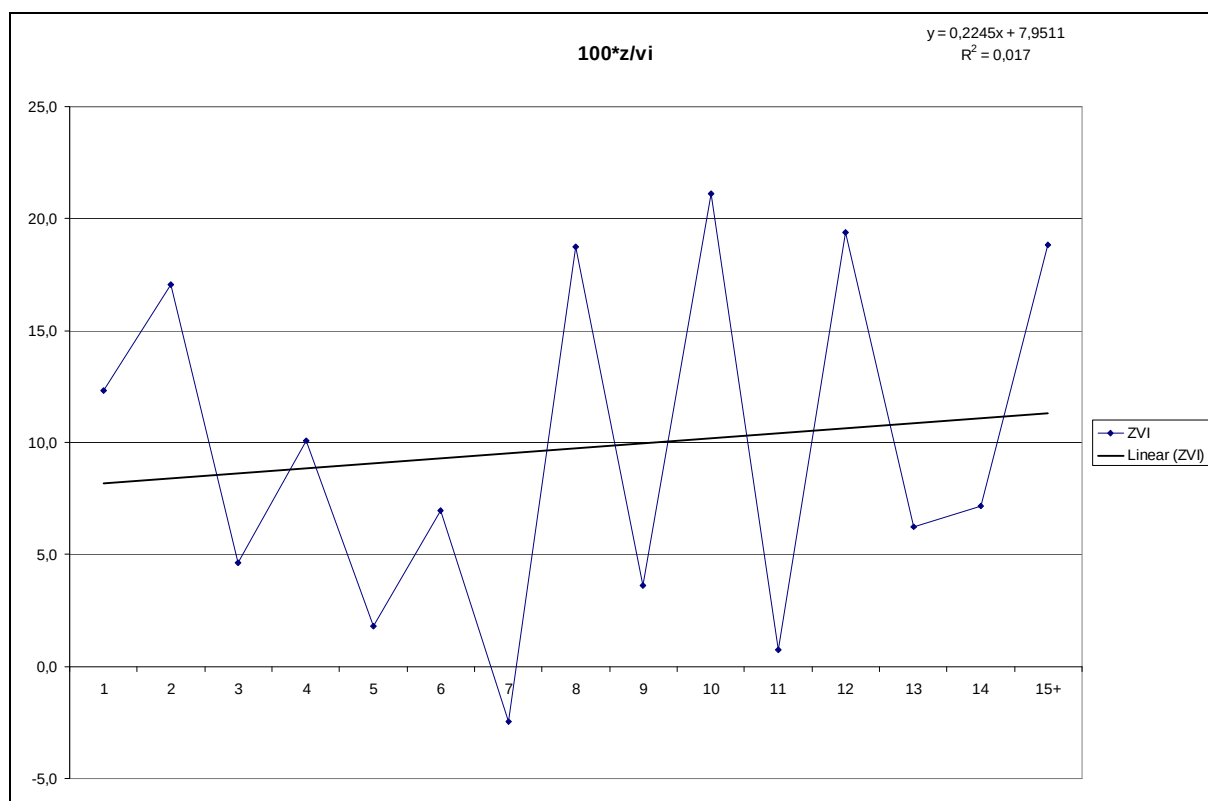
Obr. 4 Výnosy



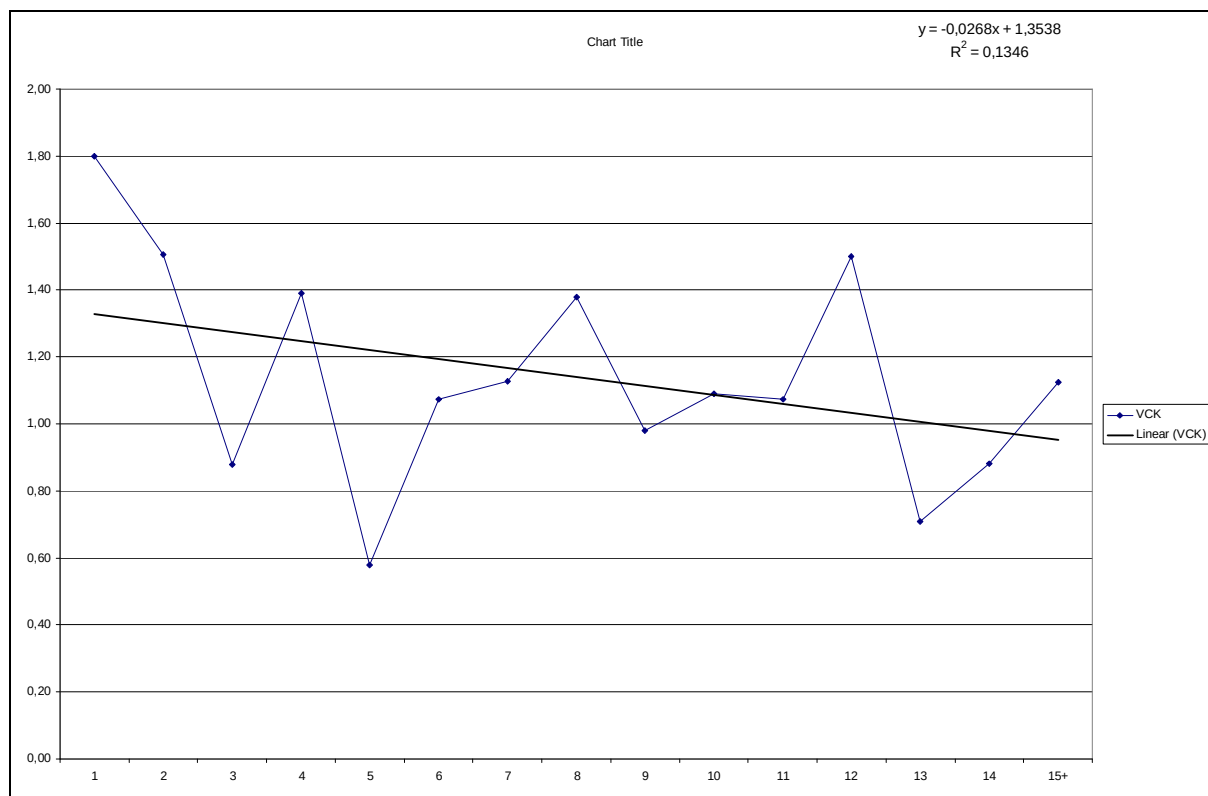
Obr. 5 Ziskovosť (rentabilita výnosov)



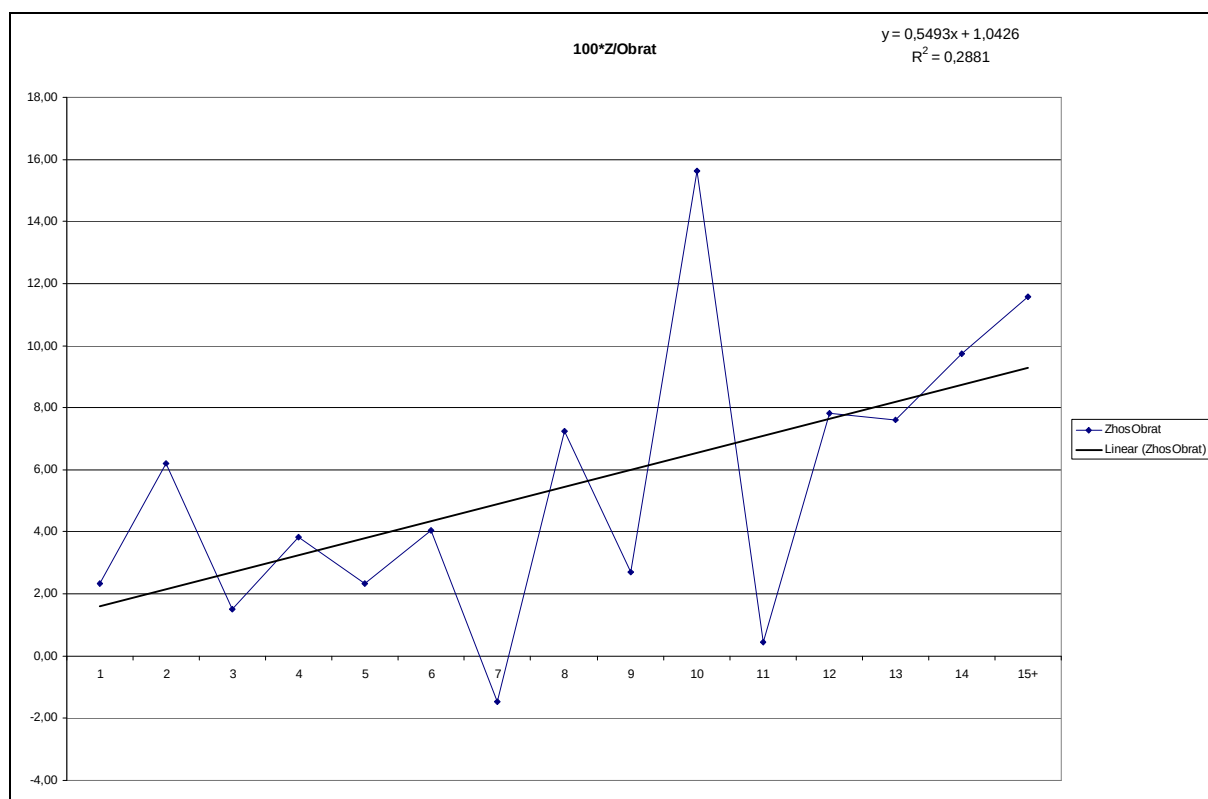
Obr. 6 Finančná produktivita práce



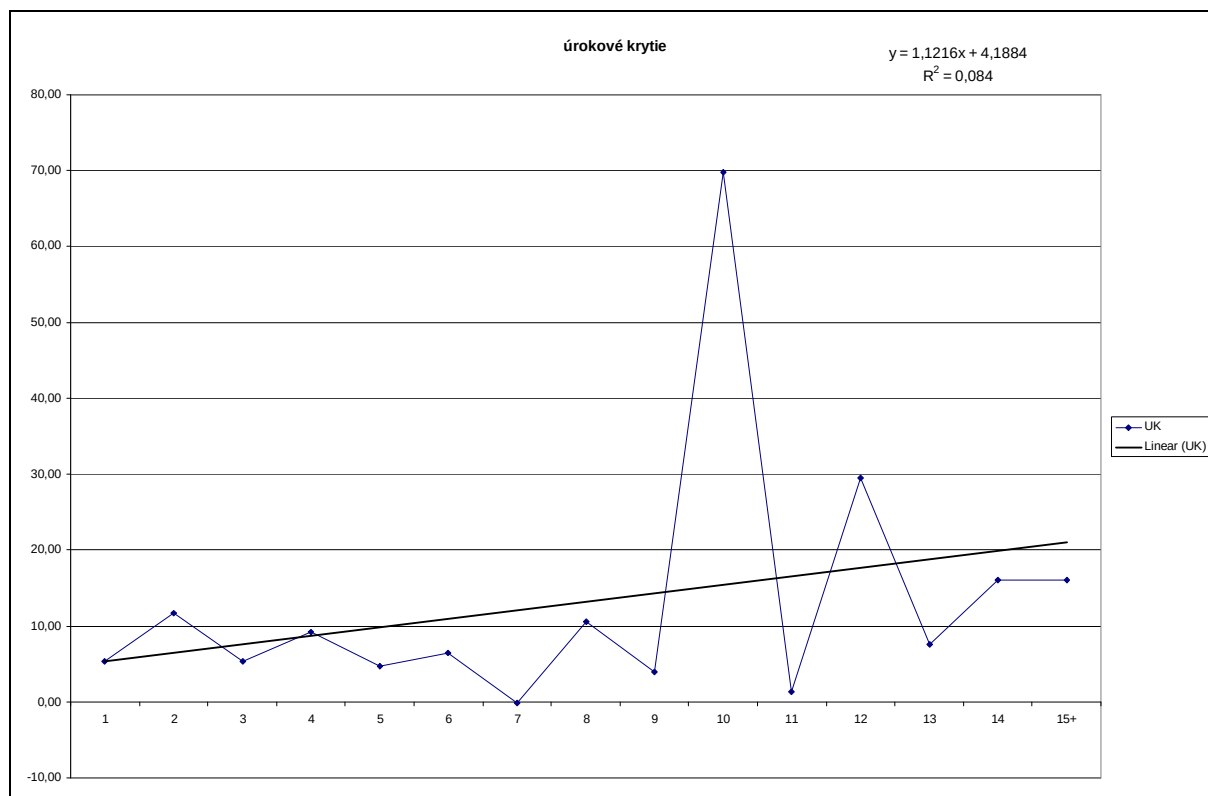
Obr.7 Rentabilita vlastného imania



Obr. 8 Účinnosť celkového kapitálu meraná výnosmi



Obr. 9 Objem zisku na obrat (rentabilita obratu)



Obr. 10 Úrokové krytie

Niekoľko poznámok k obsahu štatistiky
 Doc. Ing. Jozef Chajdiak, CSc., Statis Bratislava

Abstract: The Paper consists some note to subject of statistics.

Úlohou štatistiky je ukázať svet taký aký je! Štatistika je súčasťou procesu rozhodovania – nerozhoduje, dáva len podklady pre rozhodovanie, podporuje rozhodovanie:

- dáva číselné podklady,
- poukazuje na atypické merania a javy,
- poukazuje na závislosť resp. nezávislosť a jej mieru,
- prezentuje hromadne napozorované údaje,
- prezentuje trend a sezónnosť vývoja,
- modeluje stav alebo vývoj skúmaných objektov,
-

Štatistika nachádza svoje vyjadrenie v štatistickom skúmaní. To má:

- cieľ: podpora procesu rozhodovania,
 etapu zisťovania,
 etapu spracovania,
 etapu rozboru,
- odporúčania pre rozhodovací proces.

V procese výučby sa čiastočne vytráca zdôraznenie faktu, že štatistika je súčasťou procesu rozhodovania, jej cieľom je podpora procesu rozhodovania a výsledkom sú odporúčania pre rozhodovací proces.

V rozhodovacom procese principiálnym je potreba sa rozhodnúť. Existuje viacero schém prijatia rozhodnutia. Jedna z nich využíva štatistické podklady. Úlohou je ukázať svet taký aký je a odhadnúť, aký bude svet po realizácii prijatého rozhodnutia.

V etape zisťovania treba špecifikovať:

- množinu zisťovaných znakov (model „sveta“),
- spôsob zisťovania (úplné a neúplné zisťovanie),
- výberové zisťovanie (poznať pravdepodobnosť zaradenia každej jednotky do vyberanej vzorky),
- ostatné neúplné zisťovania.

Problematika špecifikácie množiny analyzovaných znakov sa čiastočne prenáša na iné disciplíny (na tie, ktorých sa analýza obsahovo dotýka), Na druhej strane, bez množiny zisťovaných znakov nemáme čo zisťovať, potom analyzovať a nakoniec vypracovávať podporné stanoviská pre rozhodovanie. Tak špecifikovanie množiny zisťovaných znakov je podstatnou súčasťou štatistiky.

Štatistika je súčasťou rozhodovacieho procesu. Z tohto pohľadu je dôležité, že:

- výsledky analýz (a teda aj odporúčania pre rozhodovanie) sú tak kvalitné, ako sú kvalitné zistené a spracované údaje, alebo horšie!
- nezistenie, chybné zistenie, vynechanie, zmena zisteného znamená analýzu iného sveta než v skutočnosti je!

Môžeme podrobnejšie pozeráť aj na obsah (kvalitu) analyzovaných údajov, ktoré obsahujú:

- **objektívny odraz** obsahu analyzovanej skutočnosti a chyby:

- náhodná chyba ($E[e] = 0$),
- systematická chyba (záujmy),
- nepoznanie/skreslenie časti skutočnosti (nedostatočné poznanie).

Zdá sa, že procesu zisťovania by sa mohlo pri výučbe venovať aj viac času.

V etape spracovania sa realizuje:

- zaznamenanie zistených hodnôt,
- kontrola zistených hodnôt,
- korekcie chybných resp. nejasných hodnôt,
- vytvorenie počítačového súboru údajov.

Z tohto pohľadu sa zdá potrebné presunúť časť výučby z oblasti IT (vytváranie, napĺňanie a vyberanie údajov z databáz resp. databánk) do štatistiky.

Štatistická podstata údajov:

- **každý** údaj patrí do nejakého súboru údajov,
- hodnoty v súbore údajov majú nejaké **rozdelenie**,
- rozdelenia hodnôt údajov sa modelujú **zákonom** rozdelenia pravdepodobností výskytu hodnôt (hodnota x sa vyskytuje s pravdepodobnosťou $p(x)$),
- zákon rozdelenia pravdepodobností výskytu hodnôt (hodnota x sa vyskytuje s pravdepodobnosťou $p(x)$) má silu **prírodného** zákona:
 - nevieme aká hodnota sa vyskytne v individuálnom meraní, ale v stovke meraní podiel výskytu hodnoty x bude okolo $p(x)$,
 - štatistická podstata znamená „**okolo**“ $p(x)$.

Rastom (n) počtu napozorovaných hodnôt empirické zistenia (u_n) **konvergujú** k objektívnym hodnotám (Q).

$$\lim_{n \rightarrow \infty} \Pr\{|u_n - Q| < \varepsilon\} = 1$$

Vzťah matematiky a štatistiky:

- sú samostatné disciplíny,
- pravdepodobnosť je čisto matematická disciplína,
- matematická štatistika predstavuje aplikáciu matematiky v štatistike resp. riešenie štatistických úloh matematikou,
- zisťovanie a spracovanie údajov sú čisto štatistické disciplíny.

Niekedy máme sklony zjednodušujúco za štatistiku považovať len matematickú štatistiku.

Uvedené poznámky nie sú výkladom Ústavného súdu k problematike čo a ako robiť v štatistike. Sú len autorovým odporúčaním pri rozhodovaní sa k niektorých aspektom práce so štatistikou. „Čo a ako“ sa musí rozhodnúť štatistik.

Kontakt na autora:
chajdiak@statis.biz

O chybách strednodobých prognóz vývoja ekonomiky Slovenska*

Martin Lukáčik – Adriana Lukáčiková – Karol Szomolányi

Abstract

The forecasting errors of models in predictions of last year lead us to the thought about the reasons of this fact. One problem is a strong influence of predictions made by technical time-series analysis. The second reason is an ignorance of the important signals which could be distorting for some methods used by modellers. And the last consideration is due to some skepticism of the origin results of models (which are not published completely).

Úvod

Prognózy strednodobého vývoja ekonomiky Slovenska z minulého roku ukázali, akých výrazných chýb sa môžeme dopustiť pri ekonometrickom modelovaní s cieľom konštrukcie prognóz dokonca aj na relatívne krátke obdobie. Pri tomto tvrdení vychádzame z porovnania hodnôt, ktoré boli publikované Výborom pre makroekonomické prognózy a Inštitútom finančnej politiky MF SR vo februári 2006 (pozri Tabuľku 1) a hodnôt, ktoré boli publikované vo februári 2007 (pozri Tabuľku 2) a ktoré už zohľadnili skutočné údaje z roku 2006 publikované Štatistickým úradom.

Tabuľka 1: Priemerná prognóza vývoja ekonomiky SR – VpMP a IFP MF SR – január 2006

Ukazovateľ	m.j.	Skut./Odhad		Prognóza									
		2005		2006		2007		2008		2009*			
		Výbor	IFP	Výbor	IFP	Výbor	IFP	Výbor	IFP	Výbor	IFP	Výbor	IFP
HDP; v bežných cenách	mld.Sk	1437,5	1437,6	1563,4	1574,1	1697,7	1702,5	1825,2	1821,8	1959,4	1946,7		
HDP; reálny rast	%	5,6	5,6	6,0	5,8	6,5	6,4	5,3	5,4	5,1	5,1		
Konečná spotreba domácností; reálny rast	%	5,7	5,8	5,1	4,5	5,2	4,8	4,9	4,4	4,7	4,2		
Konečná spotreba domácností; nominálny rast	%	8,8	9,0	8,4	8,1	7,6	6,8	7,2	6,5	7,2	6,5		
Priem. mes. mzda - hospodárstvo (nom. rast)	%	9,0	9,3	7,2	7,2	6,6	6,0	6,1	6,0	6,0	6,2		
Priem. mes. mzda - hospodárstvo (reálny rast)	%	6,3	6,4	3,6	3,6	4,4	4,0	3,8	4,0	3,5	3,7		
Priemerný rast zamestnanosti, VZPS	%	1,8	1,8	1,0	1,1	1,0	1,0	0,8	0,8	0,7	0,8		
Priemerný rast zamestnanosti, podľa ev. poč.	%	1,9	1,9	1,0	1,1	1,1	1,0	0,8	0,8	0,7	0,8		
Priemerná miera nezamestnanosti, VZPS	%	16,3	16,3	15,6	15,8	15,1	15,6	14,7	15,3	14,2	14,8		
Index spotrebiteľských cien (priemerný rast)	%	2,7	2,7	3,5	3,5	2,2	2,0	2,2	2,0	2,3	2,4		
Index výrobných cien (priemerný rast)	%	4,6	4,8	4,8	4,9	2,6	1,5	2,2	1,5	2,2	1,7		
Bilancia bežného účtu (podiel na HDP)	%	-5,3	-6,1	-4,2	-4,2	-2,2	-2,4	-1,5	-1,7	-0,8	-1,2		
Konečná spotreba verejnej správy (reálny rast)	%	1,6	1,6	3,2	3,3	2,6	2,9	2,6	3,0	2,9	2,9		
Tvorba hrubého fixného kapitálu (reálny rast)	%	11,1	11,5	8,4	6,5	6,6	6,0	6,2	4,0	5,1	4,0		
Export tovarov a služieb (reálny rast)	%	10,4	10,2	11,2	11,7	13,9	13,6	9,3	6,0	8,5	6,2		
Import tovarov a služieb (reálny rast)	%	9,3	9,3	10,5	10,7	11,8	11,7	8,8	5,2	8,1	5,4		

* rok 2009 – výbor bez NBS

Zdroj: Výbor pre makroekonomické prognózy

Výrazné rozdiely môžeme pozorovať najmä pri prognózach reálneho rastu exportu a importu tovarov a služieb, hodnotovo menšie ale stále dosť výrazné aj pri prognóze reálneho rastu konečnej spotreby domácností či hrubého domáceho produktu. Pri nominálnych ukazovateľoch zlyhala prognóza najsledovanejšej hodnoty HDP.

* Príspevok vznikol v rámci projektu VEGA 1/4652/07 – Analýza aktuálnych problémov vývoja slovenskej ekonomiky pred vstupom do Európskej menovej únie - ekonometrický prístup

Podobne výrazne odlišná bola prognózovaná hodnota konečnej spotreby domácností, značne podcenený bol rast zamestnanosti podľa výberového zisťovania, indexu výrobných cien alebo bilancie bežného účtu. Teda pre viac ako polovicu sledovaných ukazovateľov došlo pri prognóze najbližšieho obdobia ku chybe od dvoch až do deväť percent.

Tabuľka 2: Priemerná prognóza vývoja ekonomiky SR – VpMP a IFP MF SR – január 2007

Ukazovateľ	m.j.	Skutočnosť		Odhad		Prognóza					
		2005		2006		2007		2008		2009	
		Výbor	IFP	Výbor	IFP	Výbor	IFP	Výbor	IFP	Výbor	IFP
HDP; v bežných cenách	mld.Sk	1471,1	1471,1	1635,2	1636,8	1806,4	1809,4	1959,9	1941,0	2120,6	2074,0
HDP; reálny rast	%	6.0	6.0	7.7	7.9	8.1	8.1	6.3	5.6	5.4	5.1
Konečná spotreba domácností; reálny rast	%	7.2	7.2	6.4	6.4	6.2	6.0	5.4	4.6	5.1	4.2
Konečná spotreba domácností; nominálny rast*	%	9.8	9.8	11.4	11.7	8.8	8.2	7.6	6.7	7.7	6.8
Priem. mes. mzda - hospodárstvo (nom. rast)	%	9.2	9.2	7.9	8.1	7.3	6.6	6.3	6.0	6.5	6.2
Priem. mes. mzda - hospodárstvo (reálny rast)	%	6.3	6.3	3.3	3.5	4.5	4.3	4.0	4.0	3.8	3.7
Priemerný rast zamestnanosti, VZPS	%	2.1	2.1	3.7	3.7	1.7	1.6	1.1	0.8	0.8	0.8
Priemerný rast zamestnanosti, podľa ev. poč.	%	2.2	2.2	2.7	2.3	1.4	1.6	0.9	0.8	0.8	0.8
Priemerná miera nezamestnanosti, VZPS	%	16.2	16.2	13.6	13.8	12.5	13.2	12.0	12.9	11.6	12.4
Index spotrebiteľských cien (priemerný rast)	%	2.7	2.7	4.5	4.5	2.5	2.2	2.3	1.9	2.7	2.4
Index výrobných cien (priemerný rast)	%	4.7	4.7	8.5	8.4	3.4	3.0	2.4	2.0	2.5	2.0
Bilancia bežného účtu (podiel na HDP)	%	-8.6	-8.6	-8.0	-8.2	-4.5	-4.1	-3.1	-2.7	-2.4	-2.0
Konečná spotreba verejnej správy (reálny rast)	%	-0.6	-0.6	4.5	3.8	3.3	2.0	3.0	3.0	2.8	2.8
Tvorba hrubého fixného kapitálu (reálny rast)	%	17.5	17.5	7.6	7.0	7.0	6.2	6.3	5.0	5.7	5.0
Export tovarov a služieb (reálny rast)	%	13.8	13.8	19.8	20.8	16.7	18.9	11.2	10.0	8.1	8.8
Import tovarov a služieb (reálny rast)	%	16.6	16.6	17.6	18.8	13.6	14.2	10.0	8.7	7.2	8.0

* vrátane neziskových inštitúcií slúžiacich domácnostiam

Zdroj: Výbor pre makroekonomické prognózy

V tomto príspevku naznačíme ako takéto odchyľky môžu vzniknúť a rozoberieme aké teoretické riešenia existujú pre zlepšenie tohto stavu.

1. Technické prognózy a ich vplyv

Základným porovnávacím etalónom ekonometrických prognóz sú technické prognózy. Preto budeme sledovať, či neboli primárnym zdrojom rozhodnutia analytikov v prospech odchylených hodnôt. Analýzu čiastočne skreslí fakt, že prvotne publikované hodnoty odhadu ŠÚ, ktoré boli vtedy k dispozícii nepoužívame (nevieme z akých hodnôt vychádzali pracovníci uvádzaných inštitúcií), ale v našich výpočtoch aplikujeme súčasne uvádzané spresnené hodnoty ukazovateľov. Zameriame sa na rad HDP v bežných cenách a reálny rast HDP, KSD, Exp a Imp.

Využitím dvoch rôznych postupov technickej analýzy a to Holtovho-Wintersovho sezónneho vyrovnávania a dekompozície radov na trend a sezónnosť spolu s Boxovou-Jenkinsovou metodológiou získame hodnoty (pozri Tabuľku 3), ktoré potvrdzujú našu domnienku, že analytici majú tendenciu korigovať svoje prognózy podľa výsledkov technických prognóz. Alebo inými slovami technické prognózy ponúkli podobne podhodnotené výsledky prognóz uvedených radov ako hodnoty predkladané inštitúciami.

Tabuľka 3: Technické prognózy z údajov známych na konci roka 2005

Holt-Winters	2006	2007	2008	2009	Trend, sezóna a Box-Jenkins	2006	2007	2008	2009
HDP; b.c. (mld.Sk)	1598.9	1725.1	1851.3	1977.4	HDP; b.c. (mld.Sk)	1617.5	1772.7	1943.8	2131.3
HDP; reálny rast (%)	5.4	5.2	4.9	4.7	HDP; reálny rast (%)	7.0	6.7	6.3	5.9
KSD; reálny rast (%)	5.9	3.4	3.3	3.2	KSD; reálny rast (%)	5.9	5.0	4.8	4.8
Export; reálny rast (%)	9.7	5.0	4.8	4.6	Export; reálny rast (%)	10.3	9.8	9.7	9.3
Import; reálny rast (%)	10.9	5.1	4.8	4.6	Import; reálny rast (%)	10.5	9.8	9.1	8.7

Výsledky Tabuľky 3 samozrejme neznamenajú, že analytici konštruujú prognózy iba pomocou extrapoláčnych metód. Pokúšame sa nimi iba vysvetliť fakt, že vplyv takýchto odhadov na analytikov je nielen vo forme prognóz exogénnych premenných, ale aj v procese akceptácie variantov ekonometrických modelov. Tomu sa v procese tvorby prognózy nedá vyhnúť, ale mal by sa ohraničiť vplyv takéhoto postupu na nevyhnutné minimum.

2. Prognózy vytvárané s využitím ekonometrických modelov

Pri tvorbe prognóz využívajúcich ekonometrické modely, vznikajú chyby prognózy z troch zdrojov. Pre predikciu premennej $\hat{y}$ vo všeobecnom modeli $y = X\beta + u$ potrebujeme predpokladané hodnoty exogénnych premenných X pre prognózované obdobie, ktoré dosadíme do odhadnutého vzťahu $\hat{y} = \tilde{X}\hat{\beta}$. Chyba prognózy $e_p = \tilde{X}(\beta - \hat{\beta}) + \tilde{u} - (X_p - \tilde{X})\hat{\beta}$ je preto tvorená tromi základnými zložkami:

1. vektorom náhodnej poruchy pre prognózované obdobie $\tilde{u}$, ktorý je vždy prítomný;
2. súčinom $\tilde{X}(\beta - \hat{\beta})$, ktorý je tým menší, čím presnejšie sme uskutočnili odhad parametrov respektíve čím presnejšie zodpovedá špecifikácia modelu reálnemu stavu;
3. súčinom $(X_p - \tilde{X})\hat{\beta}$, ktorý predstavuje chybu spôsobenú nepresnými hodnotami vysvetľujúcich premenných a ktoré sa prenášajú na prognózu vysvetľovanej premennej.

Pripísať celú chybu prognózy prvej zložke by nebolo úplne správne. Každému je jasné, že dokonalá prognóza neexistuje, ale istý stupeň presnosti by sme mali byť schopní dosiahnuť. Deväťpercentná chyba prognózy pre najbližšie obdobie je však zrejme mimo rozsah akceptácie. Takže zrejším zdrojom chýb sú druhá aj tretia zložka.

Nepresnosť prognózy spôsobená nesprávnymi predpokladanými hodnotami exogénnych premenných (nekontrolovaných modelom) by sa pri uvádzaných hodnotách exportu alebo importu dala akceptovať, keby vývoj okolitých krajín alebo hlavných obchodných partnerov SR priniesol neočakávanú akceleráciu. Ale to sa neudialo, rovnako ako nedošlo k extrémne vhodnému vývoju cien strategických surovín. Preto je zřejmé, že najväčšou časťou k chybe prognóz prispela nie úplne správna štruktúra modelov, respektíve nevhodne prijaté predpoklady, či ignorovanie významných faktov nezapadajúcich do koncepcie tvorby modelu.

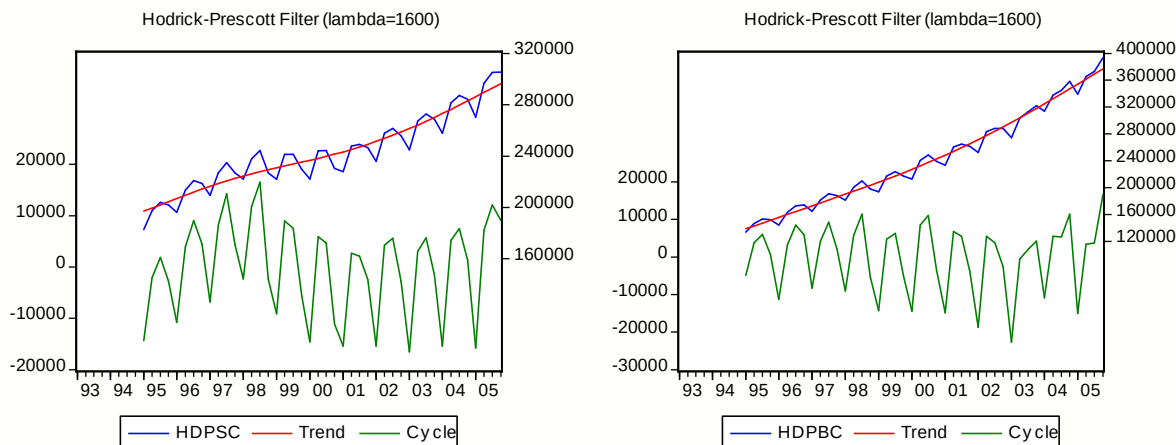
Ak akceptujeme fakt, že konštruované ekonometrické modely nedokázali identifikovať takú úroveň akcelerácie hospodárstva, ktorá nastala, tak by sme sa mali pokúsiť objaviť tie najzávažnejšie chyby, ktoré na tvorcov modelov striehnu pri ich konštrukcii. Predpokladáme pritom, že pre prognózovanie neboli využité modely všeobecnej ekonomickej rovnováhy vychádzajúce z údajov roku 2000 alebo modely vychádzajúce z ročných údajov, ktoré nemali šancu postihnúť taký výrazný dynamický akceleračný vývoj. Ten sa mohol objaviť iba v systémoch s rozvinutou štruktúrou dynamických členov – v modeloch využívajúcich diferencované nestacionárne premenné spolu s ich kointegrovanými kombináciami reprezentovanými sústavou diferenčných rovníc – modelmi ECM.

3. Problémy modelovania nestacionárnych ekonomických radov

Základným problémom modelovania ukazovateľov ekonomiky Slovenska je ich nestacionarita. Pre jej identifikáciu a určenie stupňa integrácie existujú testy jednotkového koreňa, ktoré žiaľ majú nízku silu. To znamená, že majú malú schopnosť zamietnuť nulovú hypotézu o nestacionarite, keď je to chyba. Iným nežiaducim javom je, že v prípade, ak si viaceré testy protirečia, tak analytik volí hodnotu, ktorá mu viac vyhovuje bez dodatočného skúmania, respektíve prijíma záver priamo z analýzy grafu, čo dopredu determinuje obvyklý záver o ráde integrácie jedna. Výsledkom môže byť vzniknutý problém nepravej regresie.

Druhou vážnou chybou, ktorej sa analytici dopúšťajú je chýbajúca dôsledná analýza štrukturálnych zmien. Ako príklad uveďme dva grafy HDP v bežných a stálych cenách, na ktoré sme použili Hodrickov-Prescottov filter (pozri Obrázok 1).

Obrázok 1: Vývoj HDP v stálych a bežných cenách



Z priebehu trendovej zložky v oboch prípadoch vidíme akceleráciu začínajúcu na konci roku 2001, čo potvrdia aj viaceré testy štrukturálnej zmeny v rôznych rovniciach, že posledný štvrtrok roku 2001 a prvé dva štvrtroky 2002 sú obdobia, kedy došlo k štatisticky významnej zmene parametrov modelov. Tento jav a aj ostatné ukazovatele reálnej ekonomiky poskytujúce rovnaké výsledky nás nabádajú k opatrnosti aj pri hodnotení nestacionarity. Tým myslíme nielen rozhodovanie o ráde integrácie, ale aj o type trendu.

Určenie typu trendu (deterministický alebo stochastický) v skúmaných premenných je kľúčovým pri výbere deterministickej schémy Johansenovej procedúry. S tou sa spája ďalší problém, a to je výber počtu kointegrujúcich vektorov. Idealizované riešenie výberu práve jedného kointegrujúceho vektora totiž značne uľahčí prácu s modelom, ale môže viesť k zamietnutiu správnej špecifikácie. A pri využívaní väčšieho počtu vysvetľujúcich premenných je jeden kointegrujúci vektor skôr výnimkou ako pravidlom.

Okrem uvedených rozhodovacích nástrah ekonometrického modelovania môžeme pripomenúť aj problém zotrvačnosti u analytikov. Myslíme tým zotrvanie na v minulosti vyvinutých modeloch a snahu iba o ich čo najmenšie úpravy. Tento fakt sa prezentuje obvykle ako úspešnosť už existujúceho modelu, ale môže viesť práve k neodhaleniu závažných procesov, ktoré znamenajú výraznú zmenu systému.

Záver

K zamysleniu sa nad chybami prognóz nás viedli výrazné rozdiely medzi skutočnými hodnotami a prognózami uvádzanými pre rok 2006. Z hľadiska reálneho stavu sa neočakávala taká výrazná akcelerácia zahraničného obchodu ani vývoja ekonomiky ako celku. Za tento stav prognóz je podľa nás zodpovedných niekoľko faktorov. Za jeden z najvážnejších faktorov považujeme príliš výrazný vplyv extrapoláčnych prognóz, ktoré by mali slúžiť len pre porovnanie a nie za východiskové hodnoty, ktoré následne majú napodobniť prognózy ekonometrických modelov.

Druhým faktorom by mohlo byť zjednodušovanie postupov prípravných fáz tvorby modelu, analýza stacionarity a sprievodné rozhodnutia, ktoré môžu ovplyvniť výslednú špecifikáciu modelu. Ako príklad uvedieme ignorovanie analýzy stability a neskúmanie existencie štrukturálnych zmien, ktoré môže byť zachytené ako prítomnosť jednotkového

koreňa, ale pritom v skutočnosti ide iba o jednoduché lineárne trendy s bodmi zvratu vo vývoji. Podobne nevšímavosť voči alternatívnym špecifikáciám napríklad pri viacerých kointegrujúcich vzťahoch v rámci Johansenovej procedúry môže viesť k nesprávnemu záveru.

A za poslednú príčinu považujeme prílišnú opatrnosť až bojzlivosť pri dôverovaní vytvoreným modelom, k čomu nás oprávňuje existujúca zdržanlivosť pri ich publikovaní a tým aj menšia šanca na ich prediskutovanie v širšom profesnom auditóriu ako je len kolektív tvorcov. Tým sa stráca aj možnosť kontroly a analýzy postupov ich tvorby.

Tieto postrehy nevzťahujeme na konkrétny model alebo jeho tvorcov, ale ponúkame ich ako všeobecné zamyslenie, ktoré by mohlo viesť k zlepšeniu reputácie modelov vytváraných s cieľom prognózovania ekonomického vývoja na Slovensku. Či sa zlepší úspešnosť prognóz v nasledujúcich rokoch a či analytici dokážu odhaliť čas ukončenia tejto prudkej akcelerácie ekonomiky, uvidíme už čoskoro.

Literatúra

1. Davidson, R. – MacKinnon, J. G.: *Econometric Theory and Methods*. Oxford University Press, New York 2004.
2. Dickey, D. A. – Fuller, W. A.: Distributions of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of American Statistical Association* 74, 1979, str. 427-431.
3. Greene, W. H.: *Econometric Analysis*. Prentice Hall, New Jersey 2003.
4. Johansen, S. (1995): Identifying Restrictions of Linear Equations With Applications to Simultaneous Equations and Cointegration. *Journal of Econometrics* 69, str. 111-132.
5. Lukáčik, M.: *Analýza stacionarity vybraných ekonomických indikátorov Slovenska v absolútnych hodnotách a po logaritmickej transformácii*. Zborník I. medzinárodného vedeckého semináru doktorandov, Praha 2004.
6. Lukáčik, M.: *Metodológia a stratégie ekonometrickej analýzy*. Dizertačná práca, Bratislava 2005.
7. Lukáčik, M. – Lukáčiková, A. – Szomolányi, K.: *Prognóza vývoja ekonomiky Slovenska na báze malých modelov*. 20. škola štatistiky EKOMSTAT 2006, Trenčianske Teplice.
8. Quantitative Micro Software (2005): *EViews 5.1 User's Guide*. Program Reference.

Kontakt

Ing. Martin Lukáčik, PhD., Ing. Adriana Lukáčiková, PhD., Ing. Karol Szomolányi, PhD.
 Katedra operačného výskumu a ekonometrie
 Fakulta hospodárskej informatiky
 Ekonomická univerzita v Bratislave
 Dolnozemska 1/b, 852 35 Bratislava
 tel.: 02/67295822

e-mail: lukacik@euba.sk, istvanik@euba.sk, szomolan@euba.sk

Cenové prognózy niektorých vybraných živočíšnych komodít

Božena Viktorínová¹

Abstract: The article focuses on the prognosis of producer prices of cattle beef and deep frozen chicken for the period from January to November 2007 with the application of spectral analysis method. The goal is to grasp of varying of empirical values around the trend, by the Fourier decomposition of time series on the sum of multiplication of periodic functions $\sin(\alpha.t)$ and $\cos(\alpha.t)$.

Key words: Time series, Prognosis, Spectral Analysis, Prices.

1. Úvod

Pri riešení úlohy sme vychádzali z mesačných časových radov cien hovädzieho mäsa a mrazených kureniec bez drobkov v Sk/kg bez DPH. Tieto ceny sú uvedené v Tab.1., 2., a 3.. Pri jatočných kravách a býkoch sme mali k dispozícii údaje od januára 2001 do februára 2007, vždy po 74 údajov. Pri mrazených kurencoch bez drobkov sme mali k dispozícii údaje od januára 1997 do decembra 2006, teda 120 údajov. Dostatočný počet údajov týchto časových radov nám umožnil použiť na prognózu cien mäsa spektrálnu analýzu.

Keďže naše časové rady nevykazujú konštantný trend, museli sme ich diferencovať, t.j. nahradiť časovými radmi prvých diferencií. Ak teda Fourierov rozklad stacionárneho časového radu má tvar:

$$y_t = a_1 \cdot \cos(\alpha_1 \cdot t) + b_1 \cdot \sin(\alpha_1 \cdot t), \quad (1)$$

kde α_1 a α_2 sú frekvencie z intervalu od 0 po π , po diferencovaní budeme označovať rozklad časového radu dy_t a predikované hodnoty diferencovaného časového radu $dy_t \text{ HAT}$.

Keďže predikované hodnoty časového radu boli aj centrovane, museli sme ich aj odcentrovať:

$$y_t \text{ HAT} = y_t \text{ HAT} + (t - T) \cdot \text{PRIEM} \quad (2)$$

kde $y_t \text{ HAT}$ je predikovaná hodnota pôvodného časového radu, T je posledný dátum ($t = 1, 2, \dots, T$) časového radu ktorý máme k dispozícii a PRIEM je priemerná hodnota zadaného časového radu.

Nakoniec rekurentný vzťah pre výpočet predikovaných hodnôt pôvodného časového radu pre $t = 76, 77, \dots, 84$ u hovädzieho dobytku a pre $t = 122, 123, \dots, 132$ u mrazených kurčiat bude nasledovný:

$$y_t \text{ HAT} = y_{t-1} \text{ HAT} + dy_t \text{ HAT} + \text{PRIEM} \quad (3)$$

2. Postup pri riešení danej problematiky

Vychádzali sme z nasledujúcich údajov časových radov:

¹ Božena Viktorínová, Ekonomická univerzita, Bratislava

Tabuľka 1. Ceny za jatočné kravy za SR, v Sk/kg j. hm. bez DPH

mesiac	2001	2002	2003	2004	2005	2006	2007
1.	58,89	48,21	52,27	44,15	56,62	57,76	56,82
2.	56,61	45,64	50,24	45,06	58,42	56,99	54,67
3.	56,43	47,76	48,76	44,15	60,62	59,26	
4.	53,32	51,30	45,29	44,23	63,78	61,35	
5.	47,20	53,59	44,49	47,07	65,23	61,48	
6.	45,30	45,71	43,51	55,05	65,92	61,92	
7.	52,84	44,58	41,46	59,18	65,59	61,46	
8.	49,93	44,38	43,83	59,04	65,45	60,98	
9.	51,28	49,45	46,14	57,10	65,09	60,33	
10.	54,15	53,38	46,18	56,82	61,80	60,62	
11.	54,31	49,03	45,94	54,98	59,32	58,64	
12.	52,62	49,74	44,40	55,02	57,72	58,65	

Zdroj: PPA – ATIS

Tabuľka 2. Ceny za jatočné býky za SR, v Sk/kg j. hm. bez DPH

mesiac	2001	2002	2003	2004	2005	2006	2007
1.	83,71	84,19	84,94	80,53	84,79	90,96	88,29
2.	83,42	84,18	85,57	80,50	87,65	92,47	86,32
3.	83,08	84,06	86,80	80,02	90,06	93,47	
4.	82,72	83,88	84,61	80,19	91,28	93,91	
5.	82,75	84,28	83,92	80,10	91,85	93,61	
6.	82,51	86,60	83,37	81,32	92,76	93,63	
7.	81,36	85,29	82,78	84,72	93,15	95,29	
8.	81,85	86,38	82,13	85,34	92,86	94,72	
9.	82,82	84,12	82,33	84,86	92,86	92,97	
10.	81,91	85,53	80,64	84,04	91,16	92,51	
11.	82,05	82,17	82,31	83,75	89,80	91,09	
12.	82,77	84,09	81,22	83,80	90,65	90,40	

Zdroj: PPA – ATIS

Tabuľka 3. Cena za mrazené kurence bez drobkov za SR v Sk/kg bez DPH

1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
52,39	54,86	53,00	47,99	64,14	65,00	53,00	59,07	54,68	52,18
51,90	55,29	51,77	48,80	64,17	60,83	53,00	58,00	53,81	51,11
51,40	56,53	49,88	51,75	67,46	61,36	52,50	55,30	53,76	50,54
50,77	56,63	49,96	53,00	67,72	57,44	49,92	53,38	53,49	50,57
50,82	57,34	49,26	53,20	67,50	55,46	51,35	52,11	53,02	49,26
51,19	58,11	49,92	53,61	67,31	55,00	52,19	52,23	54,57	50,00
51,71	58,35	50,27	54,50	67,78	55,42	55,58	53,01	56,61	50,46
52,06	58,49	50,06	56,71	69,33	55,50	58,54	54,59	57,49	50,90
52,98	58,50	50,87	58,25	70,38	55,63	59,79	54,60	57,47	51,88
54,34	58,22	50,65	62,00	70,25	55,63	61,64	55,75	56,92	52,64
54,19	56,92	49,76	63,50	68,33	54,88	59,78	54,70	55,66	52,14
54,00	53,22	48,85	65,00	66,67	54,63	58,19	54,40	53,93	51,86

Zdroj: PPA - ATIS

V prípade prognózy cien jatočného kravského mäsa sme použili pre výpočet predikovaných hodnôt diferencovaného časového radu vzťah:

$$dy_t HAT = 0,51377 * \cos(0,60250 * t) - 1,01595 * \sin(0,60250 * t) \quad (4)$$

a rekurentný vzťah pre výpočet predikovaných hodnôt pôvodného časového radu:

$$y_t HAT = y_{t-1} HAT + dy_t HAT - 0,0578082 \quad (5)$$

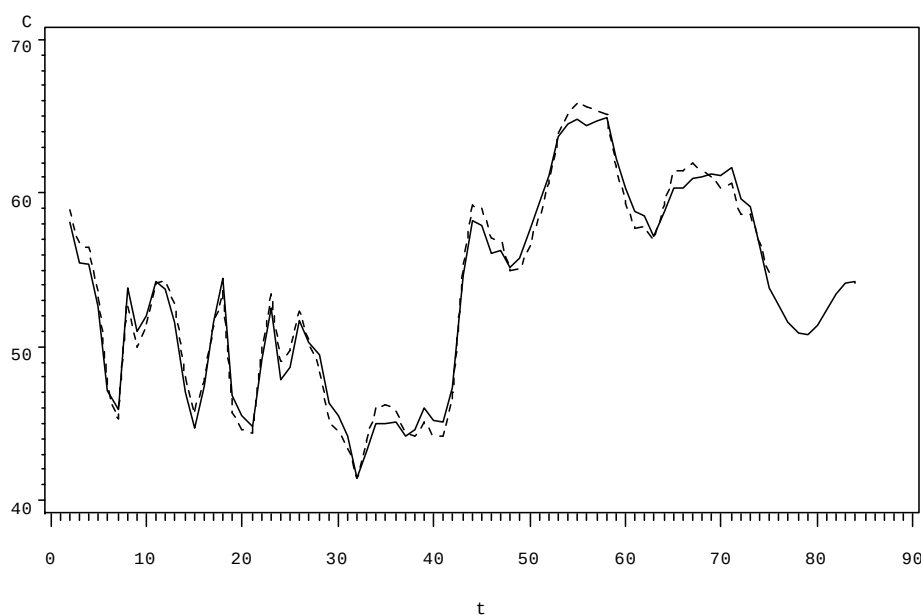
kde $t = 76, 77, \dots, 84$, $t - 1 = 75, 76, \dots, 84$. Vypočítané predikované hodnoty spolu s vypočítanými predikovanými hodnotami diferencovaného časového radu uvádzame v nasledujúcej tabuľke:

Tabuľka 4. Prognóza cien kravského mäsa v Sk/kg bez DPH

Obs.	$dy_t HAT$	y_t	$y_t HAT = \text{prognóza}$
75	- 0,76495	54,67	53,8472 febr.2007
76	- 1,10810	.	52,6813 marec
77	- 1,06102	.	51,5625 apríl
78	- 0,64030	.	50,8644
79	0,00591	.	50,8125
80	0,65003	.	51,4047
81	1,06524	.	52,4122
82	1,10532	.	53,4597
83	0,75616	.	54,1580
84	0,14072	.	54,2409 november

V poslednom stĺpci od druhého riadku Tab.4., teda od marca 2007 je uvedená prognóza cien jatočného kravského mäsa do novembra 2007.

Zhodu empirických (prerušovaná čiara) a teoretických (plná čiara) hodnôt časového radu môžeme vidieť na Obr.1.:



Obr.1 Grafické porovnanie teoretických a empirických cien kravského mäsa spolu s prognózou

ZOU

V prípade prognózy cien jatočného býčieho mäsa sme použili nasledovný vzťah pre výpočet predikovaných hodnôt diferencovaného časového radu:

$$dy_t HAT = -0,34752 * \cos(0,17214 * t) + 0,43148 * \sin(0,17214 * t) \quad (6)$$

Rekurentný vzťah pre výpočet predikovaných hodnôt pôvodného časového radu potom bude:

$$y_t HAT = y_{t-1} HAT + dy_t HAT + 0,0357534 \quad (7)$$

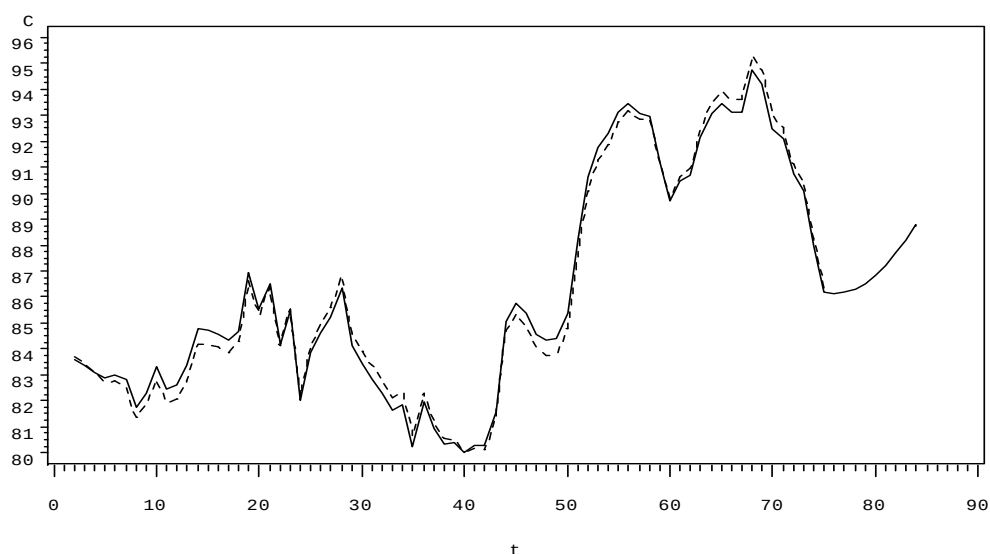
Pre $t = 76, 77, \dots, 84$ a $t - 1 = 75, 76, \dots, 84$. Vypočítané predikované hodnoty spolu s vypočítanými predikovanými hodnotami diferencovaného časového radu uvádzame v nasledujúcej tabuľke:

Tabuľka 5. Prognóza cien jatočného býčieho mäsa v Sk/kg bez DPH

Obs.	$dy_t HAT$	y_t	$y_t HAT = \text{prognóza}$
75	- 0,18149	86,32	86,1743 febr.2007
76	- 0,08914	.	86,1209 marec
77	0,00584	.	86,1625 apríl
78	0,10065	.	86,2989
79	0,19249	.	86,5271
80	0,27863	.	86,8415
81	0,35654	.	87,2338
82	0,42391	.	87,6935
83	0,47874	.	88,2080
84	0,51943	.	88,7630 november

V poslednom stĺpci od druhého riadku Tab.5., teda od marca 2007 je uvedená prognóza cien jatočných býkov do novembra 2007.

Zhodu empirických (prerušovaná čiara) a teoretických (plná čiara) hodnôt časového radu môžeme spolu s prognózou sledovať na Obr.2.:



Obr.2. Grafické porovnanie skutočných a teoretických cien jat. býčieho mäsa spolu s prognózou

V prípade prognózy cien mrazených kurčiat bez drobkov sme použili nasledovný vzťah pre výpočet predikovaných hodnôt diferencovaného časového radu:

$$dy_t = -0,80762 * \cos(0,52800 * t) - 0,13392 * \sin(0,52800 * t) \quad (8)$$

Rekurentný vzťah pre výpočet predikovaných hodnôt pôvodného časového radu potom bude:

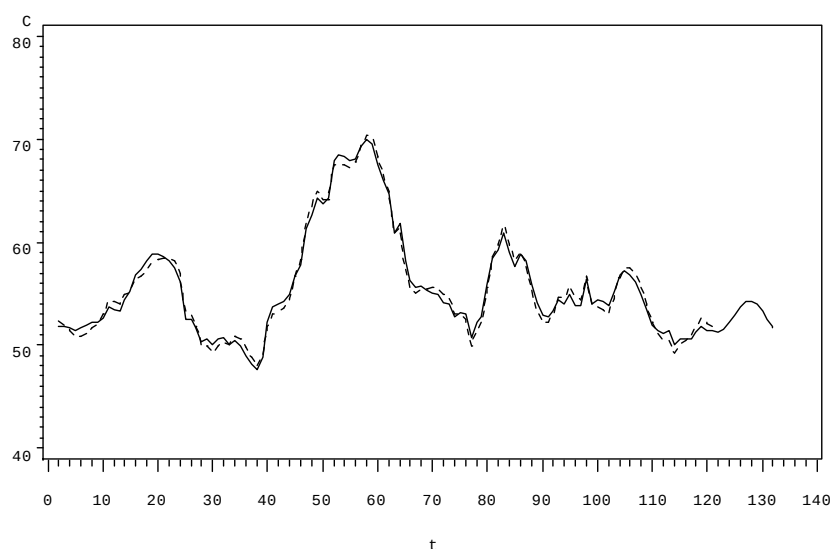
$$y_t HAT = y_{t-1} HAT + dy_t HAT - 0,0044538 \quad (9)$$

pre $t = 122, 123, \dots, 132$ a $t - 1 = 121, 122, \dots, 132$. Vypočítané predikované hodnoty spolu s vypočítanými predikovanými hodnotami diferencovaného časového radu uvádzame v nasledujúcej tabuľke:

Tabuľka 6. Prognóza cien mrazených kurčiat bez drobkov za SR v SK/kg bez DPH

Obs.	$dy_t HAT$	y_t	$y_t HAT = \text{prognóza}$
121	- 0,51421	51,86	51,3413 december 2006
122	- 0,12325	.	51,2136 január 2007
123	0,30127	.	51,5104
124	0,64374	.	52,1497
125	0,81088	.	52,9562
126	0,75716	.	53,7089
127	0,49721	.	54,2016
128	0,10185	.	54,2990
129	- 0,32126	.	53,9733
130	- 0,65687	.	53,3120
131	- 0,81356	.	52,4940
132	- 0,74868	.	51,7408 november 2007

V poslednom stĺpci Tab.6. od druhého riadku, teda od januára 2007 je uvedená prognóza cien mrazených kureniec do novembra 2007. Zhodu empirických (prerušovaná čiara) a teoretických (plná čiara) hodnôt časového radu spolu s prognózou môžeme sledovať na Or.3



Obr.3. Grafické porovnanie skutočných a teoretických cien mrazených kureniec spolu s prognózou.

3. Záver

Článok vznikol na základe potrieb praxe – vypočítať prognózy už uvedených vybraných komodít živočíšnej výroby a to v rámci grantovej úlohy VEGA 1/ 2631/ 05. Použitá spektrálna analýza sa zo všetkých vyskúšaných modelov najlepšie osvedčila na prognózovanie ,aj tým, že vykazuje veľmi dobrú zhodu medzi skutočnými a teoretickými hodnotami časových radov. Žiadna metóda ale nie je imúnna voči nepredvídateľným zmenám, napr. epidémii zvierat, čo má potom vplyv na zmenu ceny mäsa a ovplyvní výsledky prognózy.

4. Literatúra

CIPRA, T. 1986. Analýza časových řad s aplikacemi v ekonomii, Praha, SNTL, 1986
CHAJDIAK, J. – MEDVEĎOVÁ, E. 1995. Prognózovanie časových radov, Bratislava, STATIS, 1995

Adresa autora:

Mgr. Božena Viktorínová ,CSc.

Ekonomická univerzita, FHI, Dolnozemska 1, 852 35 Bratislava

viktorin@euba.sk

Impact Analysis of the New Road Law to the Accident Rate in the Czech Republic

Bohdan Linda , Jana Kubanová

Abstract

The road new law is valid in the Czech Republic from the 1st July 2006, introducing essentially harder sanctions for the drivers causing any traffic offences. The paper deals with evaluation of impact of the new law upon the traffic accidents.

Introduction

The new road law started be valid in the Czech Republic since the 1.7.2006. The tendency this law is (by implementation so-called punctual system and tighten up - substantial increasing rates- fines for traffic offences) to cause the drivers to better abidance by traffic rules. Almost one year has elapsed since introduction of this law into existence. The views of the experts and even amateurish public to efficiency of this law diametrically differ. Views from the extreme that new law has negative effect, till the second extreme, that new law brought substantial improvement of safeness in the traffic road occur in both groups. If the citizen drives on the road in the Czech Republic, he must claim, that in intravilan and extratravilan the regulated speed is kept (thanks to fear from high fines and points loss. On the other hand excessively long motorcade are formatted. In addition thanks to the insufficient high-speed road system the riding through the Czech Republic becomes one of the slower in the Europe (even in Eastern Europe). We suppose that this effect wasn't objective of the authors of the new road law. We trust that its aim was the improvement of safeness in the public communication.

Take a look at increasing road-traffic safety by an objective look of the mathematical statistics. Since even haven't elapsed all the year round since introduction of the new road law into practice, it is impossible to use annual time series containing the general indices for its efficiency evaluation. That is why the annual time series containing quarterly indices were used. That way it was possible at least in part to evaluate influence of the new law to the road-traffic safety. These time series are stated in the following table:

Tab.1.

Year	quarter	CPN	PU	PTZ	PLZ	claims
2000	Q 1	52 925	300	1 085	5 083	1 694,09
	Q 2	51 332	383	1 402	7 468	1 694,77
	Q 3	50 914	410	1 647	7 865	1 758,89
	Q 4	56 345	393	1 391	6 647	1 948,14
2001	Q 1	41 343	261	1 011	5 191	1 775,61
	Q 2	44 178	304	1 472	7 532	1 900,10
	Q 3	46 846	400	1 667	8 429	2 155,65
	Q 4	53 297	369	1 343	7 145	2 412,53
2002	Q 1	42 140	260	971	5 341	1 885,01
	Q 2	45 377	365	1 479	7 651	2 136,47
	Q 3	48 933	417	1 646	8 777	2 324,77
	Q 4	54 268	389	1 396	7 244	2 544,95
2003	Q 1	44 594	277	955	5 401	2 077,39
	Q 2	48 503	375	1 427	8 408	2 264,95
	Q 3	50 584	445	1 700	9 493	2 495,87
	Q 4	52 170	350	1 171	7 010	2 496,06
2004	Q 1	48 070	232	839	5 411	2 319,51

	Q 2	48 430	359	1 291	7 891	2 367,13
	Q 3	48 637	394	1 540	8 906	2 465,63
	Q 4	51 347	397	1 208	7 335	2 535,12
2005	Q 1	48 863	229	785	5 339	2 319,25
	Q 2	47 702	329	1 218	7 672	2 367,80
	Q 3	49 538	364	1 314	8 259	2 475,19
	Q 4	53 159	364	1 079	6 704	2 609,06
2006	Q 1	51 997	206	720	4 600	2 472,64
	Q 2	50 605	264	1 225	6 926	2 524,62
	Q 3	39 156	279	1 036	6 398	1 927,91
	Q 4	46 389	286	1 017	6329	2 199,46
2007	Q 1	39 712	209	760	4 678	1 815,95

Used shortcuts:

CPN - total number of accident damage

PU - number of deaths

PTZ - number of heavy injured

PTZ - number of light injured

The damages are expressed in the thousands of Czech Crowns (Kč)

With regard to objectively disposable data after the introduction of the new rules (3., 4. quarter 2006 and 1. quarter 2007) the statistical induction for the second quarter haven't been implemented.

Results

The estimates of the regression lines respectively regression parabola and relevant confidence intervals were calculated from the data, in chronological way preceding the beginning of the force of the new law. In the concrete, data for the first quarter in the years 2000 - 2006 were used, for the third and fourth quarter in the years 2000 - 2005. The second quarter is impossible to evaluate (since the second quarter of the year 2007 haven't finish yet). The calculations were made by the help of the statistical programme Unistat.

1/ The time series for the first quarter:

The total number of the accidents was fitted with the straight line, its equation is $y = 44535,14 + 649,5x$. The 95% confidence interval for the forecasted value is

$\langle 46071,83 ; 53390,46 \rangle$. The actual value of the total number of the accidents is 39 712.

The total number of the deaths at the traffic accidents was fitted with the straight line, its equation is $y = 305,5714 + -13,3571x$. The 95% confidence interval for the forecasted value is $\langle 187 ; 210 \rangle$. The actual value of the total number of the death is 209.

The total number of the heavy injured was fitted with the straight line

$y = 1149,286 - 59,9643x$. The 95% confidence interval for the forecasted value is $\langle 651 ; 687 \rangle$.

The actual value of the total number of the heavy injured is 760.

The total number of the light injured was fitted with the parabola

$y = 4572 + 479,8929x - 64,821x^2$. The 95% confidence interval for the forecasted value is $\langle 4163 ; 5346 \rangle$. The actual value of the total number of the light injured is 4 678.

The damages in the thousands of crowns were fitted with the straight line

$y = 1526,583 + 137,7649x$. The 95% confidence interval for the forecasted value is $\langle 2584 ; 2673 \rangle$. The actual value of the damages is 1 815,95.

2/ The time series for the third quarter:

The total number of the accidents was fitted with the straight line, its equation is $y = 49227,6 + 4,114286x$. The 95% confidence interval for the forecasted value is $\langle 47763; 50757 \rangle$. The actual value of the total number of the accidents is 39 156.

The total number of the deaths at the traffic accidents was fitted with the straight line, its equation is $y = 427 - 6,28571x$. The 95% confidence interval for the forecasted value is $\langle 352; 401 \rangle$. The actual value of the total number of the death is 279.

The total number of the heavy injured was fitted with the straight line $y = 1784,867 - 56,9143x$. The 95% confidence interval for the forecasted value is $\langle 1231; 1427 \rangle$. The actual value of the total number of the heavy injured is 1 036.

The total number of the light injured was fitted with the parabola $y = 6577,3 + 1342,0036x - 174,9107x^2$. The 95% confidence interval for the forecasted value is $\langle 7166; 9498 \rangle$. The actual value of the total number of the light injured is 6 398.

The damages in the thousands of crowns were fitted with the parabola $y = 1355,5101 + 475,4649x - 48,8113x^2$. The 95% confidence interval for the forecasted value is $\langle 2259; 2643 \rangle$. The actual value of the damages is 1 927,91.

3/ The time series for the fourth quarter:

The total number of the accidents was fitted with the straight line, its equation is $y = 55818,8 - 682,229x$. The 95% confidence interval for the forecasted value is $\langle 49836; 52250 \rangle$. The actual value of the total number of the accidents is 46 389..

The total number of the deaths at the traffic accidents was fitted with the straight line, its equation is $y = 387 - 2,85714x$. The 95% confidence interval for the forecasted value is $\langle 348; 385 \rangle$. The actual value of the total number of the death is 286.

The total number of the heavy injured was fitted with the straight line $y = 1483,667 - 62,5714x$. The 95% confidence interval for the forecasted value is $\langle 986; 1105 \rangle$. The actual value of the total number of the heavy injured is 1 017.

The total number of the light injured was fitted with the parabola $y = 6577,3000 + 1342,0036x - 174,9107x^2$. The 95% confidence interval for the forecasted value is $\langle 7166; 9498 \rangle$. The actual value of the total number of the light injured is 6 329.

The damages in the thousands of crowns were fitted with the parabola $y = 1674,3386 + 394,2450x - 41,5310x^2$. The 95% confidence interval for the forecasted value is $\langle 2042; 3047 \rangle$. The actual value of the damages is 2 199,46.

Conclusion

When we take a glance more closely at the results of the statistical evaluation we can see, that mostly real value lies either confidence intervals, or close to them. It means that the precaution stated in the new road law by far didn't bring the expected improvement. Speaking about prominent improvement in general, it is related to a number of the accident and general damages. As regards health prejudice and life, downgrade was even noted in any cases. The extremely mild winter with minimum snowfall plays its role in statistics of the 4. quarter 2006 and 1. quarter 2007. These statistics during this period can be considered as substandard. That

is why it is necessary to wait with definitive evaluation of the force of legislation. The stated results predicate that the prominent changeovers can't be expected.

Literature:

- [1] Kubanová, J.: Statistické metody pro ekonomickou a technickou praxi. Statis, Bratislava 2004.
- [2] Fahrmeier, L. and all: Statistik. Springer, Berlin 2004.
- [3] www pages of the Czech statistical office.
- [4] www pages of the Ministry of Internal Affairs of the Czech Republic.

Contacts address:

Bohdan Linda, Jana Kubanová

Studentská 95

53210 Pardubice

bohdan.linda@upce.cz , jana.kubanova@upce.cz

Autoregresní modely – řešení resamplingovou metodou

Jana Kubanová, Bohdan Linda

Universita Pardubice

Abstract:

K analýze časových řad lze použít řadu pravděpodobnostních modelů. Jejich matematická interpretace může být relativně složitá, možné zjednodušení problému nabízí bootstrapový přístup. V příspěvku jsou popsány dva modely, jedním z nich je tzv. autoregresní model prvního řádu, druhým pak autoregresní model druhého řádu.

Autoregresní model prvního řádu

Předpokládejme, že μ je střední hodnotou Y_t , kde $t = 1, 2, \dots, n$. Definujeme $Z_t = Y_t - \mu$. Střední hodnota náhodné veličiny Z_t je rovna 0. Autoregresní model prvního řádu předpokládá, že každá hodnota z_t je lineární kombinací předešlé hodnoty z_{t-1} a nezávislého poruchového členu ε_t , což můžeme vyjádřit následujícím způsobem:

$$z_t = \beta z_{t-1} + \varepsilon_t \quad (1)$$

Neznámý parametr β zde nabývá hodnot od -1 do +1.

O náhodných poruchách ε_t budeme předpokládat, že jsou náhodným výběrem z rozdělení pravděpodobností s neznámou distribuční funkcí F a střední hodnotou rovnou 0.

Dalším úkolem je získat odhad parametru β . Lze postupovat metodou nejmenších čtverců.

Prvním krokem je odhad střední hodnoty μ , kterým je $\bar{Y}$. Pak pro každé t

$$Z_t = Y_t - \bar{Y}$$

Předpokládejme, že B je odhadem skutečné hodnoty parametru β . Reziduální součet čtverců pro tento odhad je dán následujícím vztahem:

$$S_B = \sum_{t=2}^n (Z_t - B Z_{t-1})^2$$

Potom parametr β odhadujeme pomocí vztahu:

$$B = \frac{\sum_{t=2}^n z_t \cdot z_{t-1}}{\sum_{t=2}^n (z_{t-1})^2}$$

Pro střední hodnotu reziduálního součtu čtverců pak platí:

$$E(S_B) = (B - \beta)^2 E\left(\sum_{t=2}^n Z_{t-1}^2\right) + nD(\varepsilon)$$

Tato hodnota je minimální v případě, kdy hodnota veličiny B je rovna skutečné hodnotě parametru β .

Na základě dané časové řady můžeme vypočítat S_B jako funkci B a najít minimalizující hodnotu, která je odhadem parametru β .

Závažnou otázkou je, jak je tento odhad přesný. S využitím bootstrapové metody je možné určit přesnost tohoto odhadu. V popsaném modelu pracujeme se dvěma neznámými veličinami, a to neznámým parametrem β a neznámou distribuční funkcí F , tedy $Q = (\beta, F)$.

Hodnoty x jsou složeny z pozorované hodnoty y_t a odpovídající časové periody t .

Můžeme postupovat bootstrapovým algoritmem, kde $x \rightarrow \hat{Q}$, kde $Q = (\beta, F)$ je odhadnuto z původních dat. Potom je parametr β odhadnut veličinou $\hat{\beta}$, takže nám zbývá odhadnout distribuční funkci F náhodných poruch.

Parametr β je zpravidla neznámý, ale jeho odhad $\hat{\beta}$ lze využít k výpočtu přibližných hodnot poruch, tj. $\hat{\varepsilon}_t = z_t - \hat{\beta} z_{t-1}$. Odhadem distribuční funkce F je zpravidla empirická distribuční funkce $\hat{F}$ přibližných poruch.

Dále přistoupíme k řešení otázky přesnosti odhadu parametru β . Bootstrapový soubor $\hat{Q} \rightarrow x^*$ je generován v souladu se vztahem (1). Rozdíl je pouze v tom, že parametry jsou nahrazeny jejich odhady, tj. $Q = (\beta, F)$ nahrazujeme $\hat{Q} = (\hat{\beta}, \hat{F})$.

V prvním kroku vypočteme $z_1 = y_1 - \bar{y}$,

přičemž tato počáteční hodnota časové řady je považována za fixní konstantu

Bootstrapová časová řada z_t^* je počítána rekurzivně

$$\begin{aligned} z_2^* &= \hat{\beta} z_1 + \varepsilon_2^* \\ z_3^* &= \hat{\beta} z_2^* + \varepsilon_3^* \\ z_4^* &= \hat{\beta} z_3^* + \varepsilon_4^* \\ &\dots\dots\dots \\ z_n^* &= \hat{\beta} z_{n-1}^* + \varepsilon_n^* \end{aligned}$$

Bootstrapové poruchové členy ε_t^* jsou náhodným výběrem z $\hat{F}$, to znamená, že každá hodnota ε_t^* je rovna jedné z přibližných poruch s pravděpodobností $1/n$.

Bootstrapový postup opakujeme R -krát a získáme R řad náhodných poruch.

Z každé z těchto časových řad získáme bootstrapovou replikaci $\hat{\beta}^*$ odhadu $\hat{\beta}$ metodou nejmenších čtverců. Při užití autoregresního schématu prvního řádu závisí každá hodnota z_t na svém předchůdci, to je na hodnotě z_{t-1} . Tento druh závislosti je znám jako Markovův proces prvního řádu.

Autoregresní model druhého řádu

Autoregresní schéma druhého řádu rozšiřuje závislost až na člen z_{t-2} . Potom

$$z_t = \beta_1 z_{t-1} + \beta_2 z_{t-2} + \varepsilon_t \quad (2)$$

Dvojrozměrný vektor $\beta = (\beta_1, \beta_2)$ je neznámým parametrem a ε_t je poruchový člen podobně jako v předchozím modelu. Metodou nejmenších čtverců nyní získáme odhad vektoru β :

$$\hat{\beta} = (Z'Z)^{-1}Z'z,$$

$$\text{kde } z \text{ je vektor } (z_1, z_2, \dots, z_n)' \text{ a } Z \text{ je matice } \begin{pmatrix} z_{t-1} & z_{t-2} \\ z_t & z_{t-1} \\ z_{t+1} & z_t \\ \dots & \dots \\ z_{n-1} & z_{n-2} \end{pmatrix}.$$

V našem příkladě je odhad parametru $\hat{\beta}$ metodou nejmenších čtverců:

Řešíme otázku přesnosti odhadu pro autoregresní schéma druhého řádu. Liší-li se $\hat{\beta}_2$ o méně než dvě standardní chyby od nuly, interpretujeme tuto skutečnost tak, že $\hat{\beta}_2$ se významně neliší od nuly.

Závěr

Obecně lze říci, že matematické modely jsou do určité míry zjednodušením složitých reálných situací a obvykle nejsou zcela přesné. Jsou vždy kompromisním řešením mezi realitou a potřebami výzkumu. Bootstrapové metody jsou schůdným řešením v případech, kdy reálná situace vyžaduje použití složitějšího modelu, neboť matematická obtížnost modelu nesouvisí s přesností bootstrapové analýzy.

Další možnou bootstrapovou metodou aplikovanou při odhadech parametrů v časových řadách je tzv. bootstrap pohybujících se bloků. Této metodě bud věnován další článek.

Literatura:

1. EFRON, B.: The Jackknife, the Bootstrap and Other Resampling Plans. CBMS-NSF Regional Conference in Applied Mathematics, Philadelphia:SIAM 1982
2. FERNHOLZ, L.T.: Von Mises Calculus for Statistical Functionals, Lecture Notes in Statistics, Springer, New York 1983
3. KALINA, M.: Elementárne o bootstrape. In.: Sborník konference Prastan 2001, Kočovce 2001, s.94-95, ISBN 80-88946-13-1
4. LINDA, B.: Resampling Methods Accuracy in Point Estimates. In: Sborník příspěvků 15. škola statistiky EKOMSTAT'2001 – Aplikace statistických metod v hospodárskej praxi. Trenčianske Teplice 2001, s.64 – 67, ISBN 80-88946-10-7
5. RAO, R.C.: Lineární metody statistické indukce a jejich aplikace. Academia Praha 1978.

Kontaktní adresy:

Bohdan Linda, Jana Kubanová

Studentská 95

53210 Pardubice

bohdan.linda@upce.cz , jana.kubanova@upce.cz

Z histórie Ekomstatov

Jozef Chajdiak, Ján Luha

Ako sem už napísali pri príležitosti konania 20-teho Ekomstatu sme si pripomenuli nedokonalosť našej pamäte a rozhodli sa spísať aspoň chronológiu tejto akcie. Pri príležitosti 21. ročníka, pripájame ďalší dielik do skladačky prehľadu Ekomstatov.

Koncom augusta v roku 1987 sa v Liptovskom Jáne uskutočnil x-tý ročník PROBASTATu (PROBAbility and STATistics), kde sa o.i. zúčastnili aj autori tohoto príspevku. J. Chajdiak za hlavný prínos z účasti na PROBAState považuje svoju ideu zorganizovať týždňovú školu štatistiky EKOMSTAT (EKOnoMická ŠTATistika) so zameraním na aplikáciu štatistických metód a postupov na analýzu javov a procesov v sociálno-ekonomickej praxi. Diskusný kolektív na PROBAState vyberal z názvov EKONSTAT, EKOMSTAT, EKOSTAT a iných (ktoré ihneď zamietol) a nakoniec rozhodol, že bude EKOMSTAT.

Idea našla svoje praktické vyjadrenie v organizácii jarnej školy štatistiky v júni 1988 v Domove speváckeho zboru slovenských učiteľov v Trenčianskych Tepliciach. Hlavní organizátori prvého ročníka J. Chajdiak a J. Kvetko navštívili v apríli 1988 potenciálne miesto konania akcie, prezreli si Domov SZSU aj kúpeľné mesto Trenčianske Teplice a prišli k záveru, že „to je ono“.

Prvý ročník mal prísny organizačný priebeh, začínalo sa v presne určených hodinách a minútach. Prednášajúci mali obleky a kravaty, resp. šaty. Okrem denných zamestnaní prebiehali v pondelok a utorok aj večerné zamestnania do 22,00 hodiny. Časom sa formálny charakter, večerné zamestnania a čiastočne aj presnosť začiatkov vytratili a v procese výučby sa sústreďujeme najmä na obsahové otázky. Popri odbornom programe je súčasťou školy aj spoločenský program.

Postupom času sa ustálil rozvrh zamestnaní v týždňovej škole štatistiky. Prednášky odborníkov z oblasti štatistiky, ekonomiky a príbuzných disciplín na určené témy prebiehajú každý deň v pondelok až v piatok v čase od 8,00 do 12,00, s častým predĺžením až do 12,30. V pondelok, utorok a vo štvrtok, v čase od 16,00 do 17,30, prebiehajú ďalšie prednášky a krátke vystúpenia účastníkov školy.

Každý pracovný deň vo večerných hodinách prebiehajú neformálne diskusie k rôznym odborným, spoločenským a aj súkromne-životným okruhom problémov, v rámci panelových diskusií, či opekaní pri ohníčku. V prvých ročníkoch sa večerné zamestnanie „opekanie“ uskutočňovalo na peknej lúke obkolesenej lesom, na hornom okraji Trenčianskych Teplíc. Zostarnutím a spohodlnením nemenovaných účastníkov sa akcia „opekanie“ presunula na prvú terasu v areáli Domova SZSU. V stredu popoludní majú účastníci školy voľno, ktoré niekoľkokrát využili na spoločný výstup na Vapeč, vrchol s pekným širokým výhľadom na okolitú prírodu, asi 20 km od Domova SZSU, či iné výlety do okolia. Poobedňajší voľný čas časť účastníkov trávil na Zelenej žabe (kým fungovala) alebo ho využila na prechádzky po Trenčianskych Tepliciach a okolí a vzorní účastníci školy na samoštúdium (často to bola

prázdná množina). EKOMSTAT sa uskutočňuje v bezprostrednom susedstve „národnej kultúrnej pamiatky ELEKTRA“. Niektorí a niekoľkokrát sme naživo videli aj jej majiteľa.

Vynikajúca atmosféra na Ekomstate podnecuje nielen skvelé spoločenské a pracovné kontakty, ale tiež „vynachádzanie“ nových zákonov. Namenovaný spoluautor tohoto príspevku (J.L.) objavil o.i. zákon troch fernetov, zákon najkrajšieho obzerania a iné zákony, ktoré neboli na veľkú škodu zaznamenané. Pripomenieme iba zmenu úrovne významnosti z $\alpha=0.05$ na $\alpha=0.04$ (pretože v mnohých reštauráciách sa objem poldecákov scvrkol).

Prvé ročníky zabezpečoval J. Chajdiak s J. Kvetkom. Od 90-tych rokov organizovanie prevzala dvojica J. Chajdiak, J. Luha. V súčasnosti už viac rokov Programový a organizačný pracuje v zložení: J. Chajdiak – predseda, J. Luha – tajomník, S. Koróny a V. Páleník – členovia. Zborník zostavujú J. Chajdiak a J. Luha. Recenzie príspevkov: J. Chajdiak, J. Luha, S. Koróny a P. Mach.

Ekomstatu sa pravidelne zúčastňuje 20 až 30 účastníkov. V zozname účastníkov je uvedené nominálne maximum, reálny počet v Domove SZSU je o niečo nižší. Okrem účastníkov zo Slovenska, hlavne v prvých ročníkoch, to boli aj kolegovia z Pražskej katedry štatistiky VŠE a z Univerzity Pardubice, vzácna bola účasť dekana Moskovského inžinierno-štatistického inštitútu akademika Vladimíra Mchitariana.

Výborná atmosféra, odborné zameranie a družný odborný kolektív spôsobujú, že veľa účastníkov sa školy rád zúčastní viackrát. Dokumentuje to aj dotazník EKOMSTAT, ktorý už viac rokov pripravuje J. Luha. Účastníci tento dotazník vyplnia a v záverečné dni školy si so záujmom pozrú výsledky získané z ich vlastných odpovedí.

Úspešní účastníci školy štatistiky EKOMSTAT (zatiaľ, chvála Pánu Bohu alebo Matke Prírode, či intenzívnemu štúdiu, resp. samoštúdiu, t.j. všetci) získavajú Osvedčenie Slovenskej štatistickej a demografickej spoločnosti o úspešnom absolvovaní školy.

O jednotlivých ročníkoch EKOMSTATu

(1.) Jarná škola štatistiky EKOMSTAT' 88 „Využitie štatistických metód v sociálno-ekonomickej praxi“

6.-10. 6. 1988, v zozname účastníkov je zapísaných 39 osôb, z ktorých sa vyše 30 aj EKOMSTATu zúčastnilo.

Zborník zostavil J. Chajdiak, má 42 strán, obsahuje 16 odborných príspevkov.

V odbornom programe sme sa sústredili na výklad použitia systému SPSS, vybraných štatistických metód a ich aplikácií v sociálno-ekonomickej praxi a iné.

2. Jarná škola štatistiky EKOMSTAT' 89 „Využitie štatistických metód v sociálno-ekonomickej praxi“

5. – 9. 6. 1989, v zozname účastníkov je zapísaných 31 osôb.

Zborník zostavil J. Chajdiak, má 40 strán, obsahuje 13 odborných príspevkov.

V odbornom programe sme sa sústredili na výklad viacrozmerných štatistických metód a ich aplikácií v sociálno-ekonomickej praxi, prezentovaná bola tiež práca so systémom SYSTAT a iné.

3. Škola štatistiky EKOMSTAT' 90 „Využitie štatistických metód v sociálno-ekonomickej praxi“

25. – 29. 6. 1990.

Zborník zostavil J. Chajdiak, má 40 strán, obsahuje 11 odborných príspevkov.

V odbornom programe sme sa sústredili na výklad problematiky analýzy časových radov, aplikácie štatisticko-matematických metód v sociálno-ekonomickej praxi, prezentovaná bola práca so systémom SYSTAT a iné.

4. Škola štatistiky EKOMSTAT' 91 „Využitie štatistických metód v sociálno-ekonomickej praxi“

3. – 7. 6. 1991.

Zborník zostavil J. Chajdiak, má 55 strán, obsahuje 14 odborných príspevkov.

V odbornom programe sme sa sústredili na výklad práce so systémom SAS a metódy štatistickej kontroly akosti, aplikácie štatisticko-matematických metód v sociálno-ekonomickej praxi a iné.

5. Škola štatistiky EKOMSTAT' 92 „Využitie štatistických metód v sociálno-ekonomickej praxi“

1. – 5. 6. 1992, v zozname účastníkov je zapísaných 22 osôb.

Zborník zostavil J. Chajdiak, má 14 strán, obsahuje 4 odborné príspevky.

V odbornom programe sme sa sústredili na výklad plánovania experimentov, štatistické riadenie kvality a na aplikácie štatisticko-matematických metód v sociálno-ekonomickej praxi a iné.

6. Škola štatistiky EKOMSTAT' 93 „Využitie štatistických metód v sociálno-ekonomickej praxi“

30.5. – 4. 6. 1993, v zozname účastníkov je zapísaných 17 osôb.

Zborník zostavil J. Chajdiak, má 34 strán, obsahuje 9 odborných príspevkov.

V odbornom programe sme sa sústredili na výklad rodinných účtov, analýzu kvalitatívnych znakov a meranie asociácie a na aplikácie štatisticko-matematických metód v sociálno-ekonomickej praxi a iné.

7. Škola štatistiky EKOMSTAT' 94 „Využitie štatistických metód v sociálno-ekonomickej praxi“

29.5. – 3. 6. 1994.

Zborník zostavil J. Chajdiak, má 22 strán, obsahuje 6 odborných príspevkov.

V odbornom programe sme sa sústredili na výklad metód výberového skúmania, makroekonomické modely Slovenska, prácu a grafickým softvérom a na aplikácie štatisticko-matematických metód v sociálno-ekonomickej praxi a iné.

8. Škola štatistiky EKOMSTAT' 95 „Kapitálový trh – analýzy a prognózy“

5. – 9. 6. 1995, v zozname účastníkov je zapísaných 29 osôb.

Zborník zostavil J. Chajdiak, má 55 strán, obsahuje 12 odborných príspevkov. ISBN 80 – 967343 – 0 - X

V odbornom programe sme sa sústredili na výklad analýzy a prognózy na kapitálovom trhu, metódy spektrálnej analýzy, Box-Jenkonsovu metodológiu. Práca so systémom SAS, SPSS a INFINITY. Aplikácie štatisticko-matematických metód v sociálno-ekonomickej praxi a iné.

9. Škola štatistiky EKOMSTAT´ 95 „Štatistické riadenie kvality“

4.9. – 8. 9. 1995, v zozname účastníkov je zapísaných 28 osôb.

Zborník zostavil J. Chajdiak, má 58 strán, obsahuje 9 odborných príspevkov. ISBN 80 – 967343 – 1 – 8.

V odbornom programe sme sa sústredili na výklad štatistického riadenia kvality, súvisiaca práca so systémom SAS a na aplikácie štatisticko-matematických metód v sociálno-ekonomickej praxi a iné.

10. Škola štatistiky EKOMSTAT´ 96 „Finančno – ekonomické analýzy“

3.6. – 7. 6. 1996, v zozname účastníkov je zapísaných 21 osôb.

Zborník zostavil J. Chajdiak, má 68 strán, obsahuje 6 odborných príspevkov. ISBN 80 – 967343 – 5 – 0.

V odbornom programe sme sa sústredili na výklad štatistických metód analýzy finančných ukazovateľov, indexy pre ordinálne znaky a na aplikácie štatisticko-matematických metód v sociálno-ekonomickej praxi a iné.

11. Škola štatistiky EKOMSTAT´ 97 „Aplikácia štatistických metód v transformujúcej sa ekonomike“

2.6. – 6. 6. 1997, v zozname účastníkov je zapísaných 26 osôb.

Zborník zostavil J. Chajdiak, má 120 strán, obsahuje 11 odborných príspevkov. ISBN 80 – 967658 – 1 – 7. Recenzenti: J. Chajdiak, J. Luha, S. Koróny.

V odbornom programe sme sa sústredili na výklad štatistickú reguláciu výrobného procesu, ekonomické prognózy Slovenska, bayesovské odhady v poisťovníctve, štatistické a metodologické aspekty tvorby a analýzy dotazníkov a na aplikácie štatisticko-matematických metód v sociálno-ekonomickej praxi a iné.

12. Škola štatistiky EKOMSTAT´ 98 „Aplikácia štatistických metód v transformujúcej sa ekonomike“

1.6. – 5. 6. 1998, v zozname účastníkov je zapísaných 23 osôb.

Zborník zostavil J. Chajdiak, má 136 strán, obsahuje 12 odborných príspevkov. ISBN 80 – 967658 – 2 – 5. Recenzenti: J. Chajdiak, J. Luha, S. Koróny.

V odbornom programe sme sa sústredili na výklad korešpondenčnej analýzy, megatrendov v ekonomike, štatistické skúmanie analýzy kvalitatívnych znakov, rozdelenia početnosti vybraných ukazovateľov priemyslu, štatistickú reguláciu výrobného procesu, a na aplikácie štatisticko-matematických metód v sociálno-ekonomickej praxi a iné.

13. Škola štatistiky EKOMSTAT´ 99 „Aplikácie štatistických metód v hospodárskej praxi“

31.5. – 4. 6. 1999, v zozname účastníkov je zapísaných 18 osôb.

Zborník zostavil J. Chajdiak, má 132 strán, obsahuje 13 odborných príspevkov. ISBN 80 – 967658 – 9 – 2. Recenzenti: J. Chajdiak, J. Luha, S. Koróny.

V odbornom programe sme sa sústredili na výklad analýzy súboru nameraných hodnôt, exporačnú analýzu, analýzu štruktúry ekonomiky, výberové štatistické zisťovania, konjunkturálne prieskumy a na aplikácie štatisticko-matematických metód v sociálno-ekonomickej praxi a iné.

14. Škola štatistiky EKOMSTAT´ 2000 „Aplikácie štatistických metód v hospodárskej praxi“

5.6. – 9. 6. 2000, v zozname účastníkov je zapísaných 17 osôb.

Zborník zostavili J. Chajdiak a J. Luha, má 196 strán, obsahuje 21 odborných príspevkov. ISBN 80 – 88946 – 04 – 2. Recenzenti: J. Chajdiak, J. Luha, S. Koróny. Programový výbor: J. Chajdiak – predseda, J. Luha – tajomník, V. Bakošová a S. Koróny – členovia.

V odbornom programe sme sa sústredili na výklad modelovanie príčinnej závislosti, logistické regresné modely, modelovania makroekonomického vývoja SR, analýza nominálnych a ordinálnych znakov a na aplikácie štatisticko-matematických metód v sociálno-ekonomickej praxi a iné.

15. Škola štatistiky EKOMSTAT' 2001 „Aplikácie štatistických metód v hospodárskej praxi“

4.6. – 8. 6. 2001, v zozname účastníkov je zapísaných 16 osôb.

Zborník zostavili J. Chajdiak a J. Luha, má 162 strán, obsahuje 14 odborných príspevkov. ISBN 80 – 88946 – 10 – 7. Recenzenti: J. Chajdiak, J. Luha, S. Koróny. Programový a organizačný výbor: J. Chajdiak – predseda, J. Luha – tajomník, V. Bakošová, B. Linda a V. Páleník – členovia.

V odbornom programe sme sa sústredili na výklad teórie rozhodovania, analýzu rentability, makroekonomický vývoj SR, štatistické metódy výskumu verejnej mienky, rozdelenia podnikov podľa hodnôt vybraných ukazovateľov a na aplikácie štatisticko-matematických metód v sociálno-ekonomickej praxi a iné.

16. Škola štatistiky EKOMSTAT' 2002 „Štatistické metódy v marketingovom výskume I.“

2.6. – 7. 6. 2002, v zozname účastníkov je zapísaných 20 osôb.

Zborník zostavili J. Chajdiak a J. Luha, má 167 strán, obsahuje 14 odborných príspevkov. ISBN 80 – 88946 – 18 – 2. Recenzenti: J. Chajdiak, J. Luha, S. Koróny. Programový a organizačný výbor: J. Chajdiak – predseda, J. Luha – tajomník, V. Bakošová, S. Koróny a V. Páleník – členovia.

V odbornom programe sme sa sústredili na výklad korešpondenčnej analýzy, viackriteriálne hodnotenie, dotazníkový výskum – metódy organizácie a metódy analýzy dát, analýzu makroekonomického vývoja a na aplikácie štatisticko-matematických metód v sociálno-ekonomickej praxi a iné.

17. Škola štatistiky EKOMSTAT' 2003 „Štatistické metódy v marketingovom výskume II.“

1.6. – 6. 6. 2003, v zozname účastníkov je zapísaných 25 osôb.

Zborník zostavili J. Chajdiak a J. Luha, má 182 strán, obsahuje 18 odborných príspevkov. ISBN 80 – 88946 – 27 – 1. Recenzenti: J. Chajdiak, J. Luha, S. Koróny. Programový a organizačný výbor: J. Chajdiak – predseda, J. Luha – tajomník, S. Koróny a V. Páleník – členovia.

V odbornom programe sme sa sústredili na výklad skúmania súboru kvalitatívnych dát, zhlukovú analýzu, klasifikačnú analýzu, analýzu makroekonomického vývoja a na aplikácie štatisticko-matematických metód v sociálno-ekonomickej praxi a iné.

18. Škola štatistiky EKOMSTAT' 2004 „Štatistické metódy v praxi.“

23.5. – 28. 5. 2004, v zozname účastníkov je zapísaných 29 osôb.

Zborník zostavili J. Chajdiak a J. Luha, má 139 strán, obsahuje 17 odborných príspevkov. ISBN 80 – 88946 – 35 – 2. Recenzenti: J. Chajdiak, J. Luha, S. Koróny. Programový a organizačný výbor: J. Chajdiak – predseda, J. Luha – tajomník, S. Koróny a V. Páleník – členovia.

V odbornom programe sme sa sústredili na výklad merania úrovne objektov charakterizovaných ordinálnymi znakmi, metód výberového skúmania, ANOVA, štatistické riadenia kvality, analýzu makroekonomického vývoja a na aplikácie štatisticko-matematických metód v sociálno-ekonomickej praxi a iné.

19. Škola štatistiky EKOMSTAT 2005 „Štatistické metódy v praxi.“

22.5. – 27. 5. 2005, v zozname účastníkov je zapísaných 18 osôb.

Zborník zostavili J. Chajdiak a J. Luha, má 135 strán, obsahuje 16 odborných príspevkov. ISBN 80 – 88946 – 43 – 3. Recenzenti: J. Chajdiak, J. Luha, S. Koróny. Programový a organizačný výbor: J. Chajdiak – predseda, J. Luha – tajomník, S. Koróny a V. Páleník – členovia.

V odbornom programe sme sa sústredili na výklad viacrozmerných štatistických metód analýzy kvalitatívnych znakov, diskriminačnú analýzu, lineárny regresný model, použitie Box-Jenkonsovej metodológie, analýzu makroekonomického vývoja a na aplikácie štatisticko-matematických metód v sociálno-ekonomickej praxi a iné.

20. Škola štatistiky EKOMSTAT 2006 „Štatistické metódy v praxi.“

21.5. – 26. 5. 2006, v zozname účastníkov je zapísaných 29 osôb.

Príspevky účastníkov sú publikované v časopise SŠDS FORUM STATISTICUM SLOVACUM 2/2006, ISSN 1336-7420. Publikáciu zostavili J. Chajdiak a J. Luha, má 226 strán, obsahuje 27 odborných príspevkov. Recenzenti: J. Chajdiak, J. Luha, P. Mach, S. Koróny. Programový a organizačný výbor: J. Chajdiak – predseda, J. Luha – tajomník, S. Koróny a V. Páleník – členovia.

Vzhľadom k jubilejnému charakteru školy štatistiky tematické zameranie je širokospektrálne s cieľom poukázať na bohatosť štatistiky a jej aplikácií a iné.

21. Škola štatistiky EKOMSTAT 2007 „Štatistické metódy v praxi.“

3.6. – 8. 6. 2007, v zozname účastníkov je zapísaných 34 osôb.

Príspevky účastníkov sú publikované v časopise SŠDS FORUM STATISTICUM SLOVACUM 1/2007, ISSN 1336-7420. Publikáciu zostavili J. Chajdiak a J. Luha, má 205 strán, obsahuje 24 odborných príspevkov. Recenzenti: J. Chajdiak, J. Luha, S. Koróny. Programový a organizačný výbor: J. Chajdiak – predseda, J. Luha – tajomník, S. Koróny, V. Páleník a V. Kvetan – členovia.

V odbornom programe sme sa sústredili na výklad: metódy kvótových výberov, štatistického riadenie výroby netvarovaných materiálov v spoločnosti Refrako, analýzy meracích systémov, modelovania príčinnej závislosti viacrozmernými regresnými modelmi, logistickej regresie a jej využitie v ekonomickej praxi, metód zhlukovej analýzy a faktorovej analýzy, kvantitatívnych metód analýzy postavenia podniku na trhu, analýzy makroekonomického vývoja a aplikácií štatisticko-matematických metód v sociálno-ekonomickej praxi a iné.

Zoznam účastníkov 21. Školy štatistiky EKOMSTAT 2007, 3.-8.6.2007

P.C.	Priezvisko,meno	Pracovisko
1	Adamica Peter	Trexima, Bratislava
2	Bakošová Viera	INFIN, Bratislava
3	Bašťovanský Tomáš	Trexima, Bratislava
4	Bujňáková Zuzana	Refrako, Košice
5	Burešová Lucia	StatSoft, Praha
6	Fertaľová Jana	PU, Prešov
7	Haluška Ján	INFOSTAT, Bratislava
8	Chajdiak Jozef	Statis, Bratislava
9	Chovancová Zuzana	ZSE, Bratislava
10	Ištoňová Iveta	Sensus, Stará Turá
11	Kadlečík Ľubomír	Trexima, Bratislava
12	Kapustova Mária	Bratislava
13	Kaščáková Alena	UMB, Banská Bystrica
14	Koróny Samuel	UMB, Banská Bystrica
15	Kubanová Jana	UP, Pardubice
16	Kvetan Vladimír	EÚ SAV, Bratislava
17	Linda Bohdan	UP, Pardubice
18	Luha Ján	ÚVVM pri ŠÚ SR, Bratislava
19	Marková Jana	INFIN, Bratislava
20	Mihály Géza	Trexima, Bratislava
21	Mitřík Vlado	ELEKTROSERVIS, Košice
22	Mlynek Martin	EÚ SAV, Bratislava
23	Nedelová Gabriela	UMB, Banská Bystrica
24	Olexa Michal	INFOSTAT, Bratislava
25	Páleník Viliam	EÚ SAV, Bratislava
26	Poláková Viera	MMS, Bratislava
27	Radvanský Marek	EÚ SAV, Bratislava
28	Rusinko Peter	Refrako, Košice
29	Stankovičová Iveta	FM UK, Bratislava
30	Šoltés Erik	FHI EU, Bratislava
31	Úradníček Vladimír	UMB, Banská Bystrica
32	Vagaský Martin	INFIN, Bratislava
33	Vester Miroslav	INFIN, Bratislava
34	Vojtková Mária	FHI EU, Bratislava

OBSAH

Úvod	1
Luha J.: Kvóťový výber	2
Bujňáková Z., Rusinko P.: Aplikovaný vývoj žiarobetonu - vplyv zrnitosti a použitej suroviny na fyzikálno-chemické parametre	18
Išťoňová I.: Analýza meracích systémov. (metóda R & R)	23
Šoltés E.: Posúdenie závislosti nákladov domácností na bývanie od vybraných faktorov metódami regresnej a korelačnej analýzy	30
Stankovičová I.: Logistická regresia a jej využitie v ekonomickej praxi	42
Koróny S.: Dvojkroková zhluková analýza – automatické zistenie počtu zhlukov	55
Chajdiak J.: Makroekonomický vývoj SR do apríla 2007	67
Vojtková M.: Hľadanie skrytej štruktúry pri analýze efektívnosti priemyselných podnikov	76
Úradníček V.: Vybrané metódy určovania váh kritérií pri analyzovaní postavenia podniku na trhu	84
Kvetan V., Mlýnek M., Radvanský M., Viliam Páleník V.: Hlavné trendy vo vývoji ekonomiky SR	91
Páleník V., Mlýnek M.: Miera zamestnanosti ako štruktúrny ukazovateľ lisabonskej stratégie	100
Kvetan V., Mlýnek M., Radvanský M., Páleník V.: Konštrukcia modelov HERMIN a CGE pre hodnotenie Národného Strategického Referenčného Rámca (NSRR)	115
Kvetan V., Radvanský M.: Regionálna prognóza SR so zameraním na trh práce a HDP do roku 2017	124
Haluška J., Olexa M.: Analýza vývoja slovenskej ekonomiky v roku 2006a prognóza jej vývoja v roku 2007(s výhľadom do roku 2010)	137
Haluška J.: Rýchle odhady vývoja základných makroekonomických ukazovateľov Slovenskej ekonomiky (Metodológia a výsledky)	149
Chajdiak J.: Analýza vplyvu úrovni zvoleného faktora na variabilitu hodnôt ekonomických ukazovateľov (ANOVA) – formulácia úlohy, model, postup analýzy, príklad	153
Chajdiak J., Holíčková E.: Vývoj vybraných ukazovateľov priemyslu podľa intervalov vytvorenej pridanej hodnoty v rokoch 1995, 2000 a 2006	163
Chajdiak J., Vester M.: Vypovedacia schopnosť firemného indexu transparentnosti	170
Chajdiak J.: Niekoľko poznámok k obsahu štatistiky	178
Lukáčik M., Lukáčiková A., Szomolányi K.: O chybách strednodobých prognóz vývoja ekonomiky Slovenska	180
Viktorínová B.: Cenové prognózy niektorých vybraných živočíšnych komodít	185
Linda B., Kubanová J.: Impact Analysis of the New Road Law to the Accident Rate in the Czech Republic	191
Kubanová J., Linda B.: Autoregresní modely – řešení resamplingovou metodou	195
Chajdiak J., Luha J.: Z histórie Ekomstatov	198
Zoznam účastníkov 21. Školy štatistiky EKOMSTAT 2007, 3. – 8. 6. 2007	204

FORUM STATISTICUM SLOVACUM

Vydavateľ

Slovenská štatistická a demografická
spoločnosť
Miletičova 3
824 67 Bratislava 24
Slovenská republika

Redakcia

Miletičova 3
824 67 Bratislava 24
Slovenská republika

Fax

02/63812565

e-mail

chajdiak@statis.biz
Jan.Luha@statistics.sk

Registráciu vykonalo

Ministerstvo kultúry Slovenskej republiky

Registračné číslo

3416/2005

Tematická skupina

B1

Dátum registrácie

22. 7. 2005

Objednávky

Slovenská štatistická a demografická
spoločnosť
Miletičova 3, 824 67 Bratislava 24
Slovenská republika
IČO: 178764
Číslo účtu: 0011469672/0900

ISSN 1336-7420

Redakčná rada

RNDr. Peter Mach – *predseda*

Doc. Ing. Jozef Chajdiak, CSc. – *šéfredaktor*

RNDr. Ján Luha, CSc. – *tajomník*

členovia:

Ing. Mikuláš Cár, CSc.

Ing. Ján Cuper

Ing. Edita Holičková

Doc. RNDr. Ivan Janiga, CSc.

Ing. Anna Janusová

RNDr. PaedDr. Stanislav Katina, PhD.

Prof. RNDr. Jozef Komorník, DrSc.

RNDr. Samuel Koróny

Doc. Ing. Milan Kovačka, CSc.

Doc. RNDr. Bohdan Linda, CSc.

Prof. RNDr. Jozef Mládek, DrSc.

Doc. RNDr. Oľga Nánásiová, CSc.

Doc. RNDr. Karol Pastor, CSc.

Prof. RNDr. Rastislav Potocký, CSc.

Doc. RNDr. Viliam Páleník, PhD.

Ing. Iveta Stankovičová, PhD.

Doc. RNDr. Beata Stehlíková, CSc.

Prof. RNDr. Michal Tkáč, CSc.

Ing. Vladimír Úradníček, PhD.

Ing. Boris Vaňo

Doc. MUDr Anna Volná, CSc., MBA.

Ing. Mária Vojtková, PhD.

Prof. RNDr. Gejza Wimmer, DrSc.

Mgr. Milan Žirko

Ročník

III.

Číslo

1/2007

Cena výtlačku 500 SKK / 20 EUR

Ročné predplatné 1500 SKK / 60 EUR