



ZÁPADOČESKÁ
UNIVERZITA
V PLZNI

PREDIKCE VÝVOJE NEZAMĚSTNANOSTI V PLZEŇSKÉM KRAJI

Vzorová semestrální práce s metodickými
pokyny pro účely KIV/MRF



2007



Predikce nezaměstnanosti v Plzeňském kraji

Tento dokument byl vypracován v rámci projektu financovaného z fondu rozvoje vysokých škol. Slouží jako učební materiál, který ukázkově demonstruje postup a maximální rozsah, jaký by měla mít semestrální práce předmětu MRF.

Vzhledem k tomu, že spektrum možností zadání semestrálních prací v rámci tohoto předmětu je značně široké, student by si měl zejména povšimnout formy strukturování dokumentu, uváděných popisů dat, metodiky měření. Student by si měl uvědomit důležitost uvádění těchto postupů na výsledek.

Současně ve většině prací se studenti budou zabývat sestavováním modelů. Předkládáme jim tímto inspiraci a demonstraci možných postupů identifikace, kvantifikace, verifikace a aplikace jednotlivých modelů na reálných datech.

Postupy jsou popsány v předkládaném dokumentu *KauzálníModelNezaměstnanosti.doc*, data a reálné sestavení modelů nad nimi je uvedeno viz příkládaný Excelovský sešit *KauzálníModelNezaměstnanosti.xls*. V textu je na tento sešit odkazováno zmínkou Excel a jméno listu.

V textu jsou barevně odlišeny metodické vsuvky barevně – oranžovou barvou. Měly by studenta při vypracovávání práce nasměrovat.



Predikce nezaměstnanosti v Plzeňském kraji

Obsah

1	ZADÁNÍ	4
2	POPIS PROBLÉMU	4
2.1	TEORETICKÝ POHLED NA NEZAMĚŠTNANOST	4
2.2	POPIS ZDROJŮ DAT	6
2.3	POPIS SOUČASNÉHO STAVU A VÝCHODISEK ŘEŠENÍ	8
2.4	POPIS MOŽNÝCH VARIANT ŘEŠENÍ	10
2.5	VYBRANÁ VARIANTA ŘEŠENÍ.....	10
3	POPIS A PODROBNÁ DOKUMENTACE VYBRANÉHO ŘEŠENÍ.....	11
3.1	POPIS DAT	11
3.2	SPECIFIKACE EKONOMETRICKÉHO MODELU	12
3.3	KVANTIFIKACE EKONOMETRICKÉHO MODELU	14
3.4	VERIFIKACE MODELU	17
3.5	PRAKTICKÉ VYUŽITÍ – PROGNÓZOVÁNÍ.....	19
4	DISKUSE	21
4.1	DISKUSE NESOULADŮ MEZI TECHNICKÝM A FUNDAMENTÁLNÍM MODELEM	21
4.2	KRITIKA DAT	21



Predikce nezaměstnanosti v Plzeňském kraji

1 Zadání

Předmětem této studie je sestavení kauzálního modelu vývoje nezaměstnanosti v některých okresech Plzeňského kraje. Tento model bude popisovat toky nezaměstnaných mezi okresy. Hlavním cílem je provedení predikce počtu nezaměstnaných v jednotlivých okresech pro nejbližší 3 roky na základě sestaveného kauzálního modelu.

V této práci budou zkoumány čtyři okresy Plzeňského kraje (viz Obrázek 1):

- Plzeň Město ... PM
- Plzeň Jih ... PJ
- Plzeň Sever ... PS
- Rokycany ... R



Obrázek 1

2 Popis problému

V této části je provedena analýza samotného problému a navržena řešení.

2.1 Teoretický pohled na nezaměstnanost

Měli byste se zaměřit na to, co budete zkoumat z teoretického pohledu. Často za tímto účelem využijete znalosti z makroekonomie či praxe.

Dříve než začneme navrhovat vlastní kauzální model, připomeneme v krátkosti základní vlastnosti a vztahy, které platí pro nezaměstnanost.

Pro výpočet míry nezaměstnanosti platí následující vztah:

$$(2.1) \quad \text{míra nezaměstnanosti} = \frac{U_e}{LF},$$

kde

U_e ... (Unemployment)... počet nezaměstnaných, kteří patří do jedné z následujících skupin:

- Nemá zaměstnání a v průběhu posledních 4 týdnů aktivně hledal práci
- Čeká na povolání zpět do zaměstnání
- Do nového zaměstnání se ohlásí v průběhu 4 týdnů

LF ... (Labor Force) = zaměstnaní + U_e ... velikost pracovní síly na trhu práce

Problematika zjišťování tohoto indexu míry nezaměstnanosti v praxi je uvedena v odstavci 2.2.

Podle příčiny nezaměstnanosti rozlišujeme následující **typy nezaměstnanosti**:

- **Frikční (dočasná)** – je způsobena přirozenou fluktuací na trhu práce (absolventi, vracející se ženy z mateřské, změna pracovního místa)



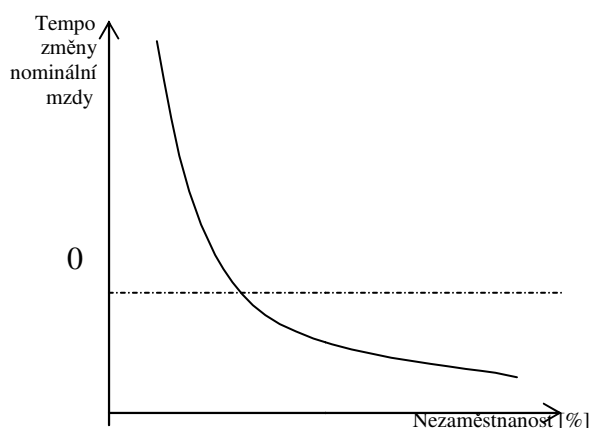
Predikce nezaměstnanosti v Plzeňském kraji

- **Strukturální** – její příčinou je změna struktury dané ekonomiky (mění se struktura průmyslových odvětví)
- **Cyklická** – souvisí s fází hospodářského cyklu, ve které se země nachází (propouštění v období recese)

Další typy nezaměstnanosti podle jiných kritérií:

- **Sezónní** – souvisí s ročním obdobím
- **Dobrovolná** – osoba setrvává dobrovolně nezaměstnaná, nehledá práci
- **Dlouhodobá** – nezaměstnaný nemá práci déle než rok

Důležitou vlastností nezaměstnanosti je vzájemná závislost nezaměstnanosti a inflace, přesněji řečeno, mezi nezaměstnaností a inflací existuje substituční efekt. Pod pojmem inflace si v tomto případě představujeme inflační očekávání či mzdovou inflaci, popsanou tempem růstu nominálních mezd. Tento nepřímý vztah popisuje Phillipsova křivka – viz Obrázek 2.



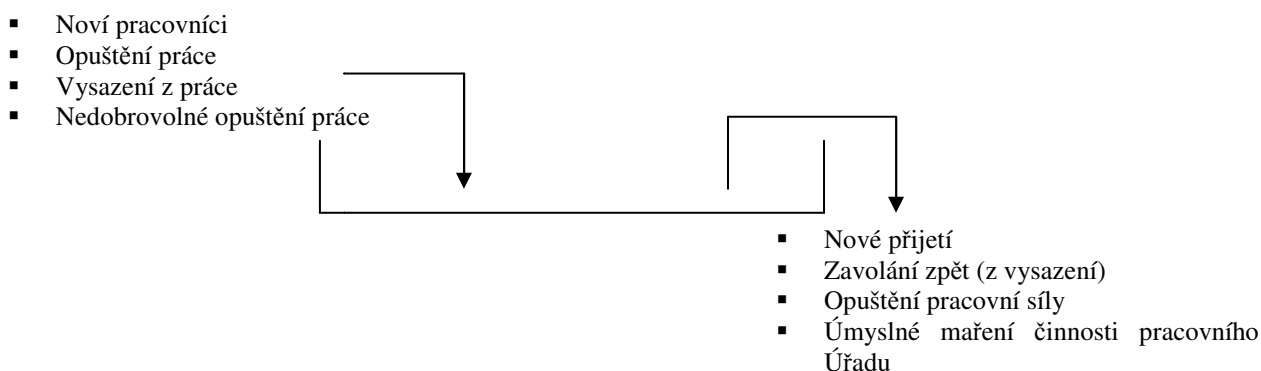
$$g_w = \frac{w_t - w_{t-1}}{w_{t-1}}$$

g_w ... tempo změny nominální mzdy

w_t ... nominální mzda v čase t

Obrázek 2

Kauzální model, který bude použit k provedení predikce je založen na tocích nezaměstnaných mezi okresy. Následující Obrázek 3 popisuje toky nezaměstnaných v rámci jednoho pracovního úřadu. (Může však také jít o toky v rámci jednoho okresu.)



Obrázek 3

Z obrázku je patrné, že nezaměstnanost roste v případě, že „přítok“ je větší než „odtok“. Důvody „přítoku“ jsou následující.



Predikce nezaměstnanosti v Plzeňském kraji

Důvody nezaměstnanosti:

- *Nově příchozí* - hledá práci poprvé
- vrací se znovu k hledání práce po více jak 4 týdnech
- *Opuštění práce* – odejde z práce a zaregistruje se jako nezaměstnaný
- *Vysazení z práce* – zastavení výplaty „bez předsudků vůči pracovníkovi“ po dobu delší než 7 dnů
- *Nedobrovolné opuštění práce* – bude propuštěn či firma zastavuje svou činnost

Kauzální model, který bude výchozí pro provedení predikce bude založen na faktu, že v případě těsně sousedících okresů dochází k situaci, kdy část „odtoků“ nezaměstnaných bude směřovat do jiného okresu či okresů (osoby získají zaměstnání v jiném okresu a za prací dojíždějí. Tím snižují v tamním okrese nabídku pracovních míst a ve svém domácím okrese nezaměstnanost). Zároveň naopak přítok nezaměstnaných může být způsoben ztrátou zaměstnání v jiném okrese.

2.2 Popis zdrojů dat

Zaměřte se na to, jak funguje měření dat spojených se zkoumanou problematikou v praxi. Nejdříve identifikujte zdroje takových dat a uveďte na ně link. Poté popište a identifikujte sledované ukazatele, co vše zahrnují a popište metodiku jejich měření. Současně je také vhodné zamyslet se, kdy a jak často jsou data měřena a pro jaký časový interval nás data budou zajímat.

Po krátkém teoretickém úvodu se podíváme na to, jak je v praxi měřen makroekonomický ukazatel nezaměstnanosti - viz (2.1) a počet nezaměstnaných. Informace jsou čerpány z Českého statistického úřadu – dále jen ČSÚ – viz [1].

Zaměstnaní jsou všechny osoby 15leté a starší, které během referenčního týdne příslušely do následujících kategorií:

1. *placení zaměstnaní*, kteří spadají do těchto skupin:

- *"v práci"* - osoby, které během referenčního týdne vykonávaly nějakou práci za mzdu nebo plat a odměna jim byla vyplácena v penězích nebo naturáliích. Není přitom rozhodující, zda jejich pracovní aktivita měla trvalý, dočasný, sezónní či příležitostný charakter a zda měly jen jedno nebo více souběžných zaměstnání.
- *"v zaměstnání, ale ne v práci"* - Jedná se o osoby, které již pracovaly ve svém současném zaměstnání, ale během vykazovaného období nebyly dočasně v práci a přitom měly formální vazbu k tomuto zaměstnání. Formální vazbou k zaměstnání se rozumí především pracovní poměr (v současné legislativní úpravě pracovní smlouva, jmenování a volba), dále dohoda o provedení práce a dohoda o pracovní činnosti, popř. další smluvní vztahy mimo oblast pracovního práva (např. autorské smlouvy)¹.

2. *zaměstnaní ve vlastním podniku*, kteří spadají do těchto skupin:

- *"v práci"* - osoby, které během referenčního týdne vykonávaly nějakou práci pro zisk nebo rodinný příjem bez ohledu na to, zda byl vyjádřen v penězích nebo v naturáliích.

¹ Problematika pracovního práva je popsána viz Zákoník práce – zahrnuje klasický zaměstnanecký poměr, dohodu o pracovní činnosti a dohodu o provedení práce. Další smluvní vztahy mimo oblast pracovního práva - problematika autorských děl popsána viz Zákon o právu autorském. Dále pak smlouva o dílo, nahodilý příjem – viz Občanský zákoník §10, odst. (1).



Predikce nezaměstnanosti v Plzeňském kraji

- "v podniku, ale ne v práci" - osoby s vlastním podnikem, které nebyly z jakéhokoli důvodu během referenčního týdne dočasně v práci.

Pojem práce je interpretován jako práce alespoň po dobu jedné hodiny v referenčním týdnu. Za zaměstnané jsou považováni i učni, studenti, osoby v domácnosti atd., kteří dostávají mzdu, plat nebo odměnu podle stejného principu jako jiné osoby. Do skupiny zaměstnaných naopak nejsou automaticky zahrnovány osoby na další mateřské (rodičovské) dovolené.

Nezaměstnaní jsou všechny osoby 15leté a starší, které ve sledovaném období současně splňovaly dále uvedené tři podmínky:

- nebyly zaměstnané (viz definice výše),
- hledaly aktivně práci².
- byly připraveny k nástupu do práce, tj. během referenčního období byly k dispozici okamžitě nebo nejpozději do 14 dnů pro výkon placeného zaměstnání nebo zaměstnání ve vlastním podniku.

Pokud osoby nesplňují alespoň jednu ze tří uvedených podmínek, jsou klasifikovány jako zaměstnané nebo ekonomicky neaktivní. Jedinou výjimkou je skupina osob, které práci nehledají, protože ji již našly, ale nástup je stanoven na dobu nejpozději do 14 dnů. Tyto osoby jsou podle definice Eurostatu zařazeny rovněž mezi nezaměstnané.

Klasifikace nezaměstnaných:

1. *registrovaní nezaměstnaní* - představují tu část nezaměstnaných, která pro aktivní hledání zaměstnání využívá služeb profesionálních organizací, ať již státních (úřady práce) nebo soukromých (zprostředkovatelny práce).
2. *neregistrovaní nezaměstnaní* jsou nezaměstnaní splňující obecné podmínky nezaměstnanosti, kteří aktivně hledají zaměstnání jinou formou než prostřednictvím organizací zprostředkovávajících práci, v případě České republiky prostřednictvím úřadů práce.

Relativní ukazatel nezaměstnanosti, počet nezaměstnaných

V této práci budeme používat stejnou konstrukci ukazatele nezaměstnanosti, jakou používá Ministerstvo Práce a Sociálních Věcí (MPSV). Konstrukce vychází z dostupných zdrojů úřadů práce a státní statistiky, kterými jsou

A. do 3. čtvrtletí 2004:

- *Čítatel*: přesná evidence **registrovaných neumístěných uchazečů** o zaměstnání vedená úřady práce v okrese trvalého bydliště uchazeče.
- *Jmenovatel*: **počet zaměstnaných v národním hospodářství** s jediným nebo hlavním zaměstnáním podle výsledků výběrového šetření pracovních sil – dále jen VŠPS (klouzavý průměr posledních čtyř čtvrtletí) + **počet registrovaných neumístěných uchazečů o zaměstnání evidovaných na úřadech práce** (klouzavý průměr posledních 12 měsíců).

² Formou aktivního hledání práce se rozumí hledání prostřednictvím úřadu práce nebo soukromé zprostředkovatelny práce, dále hledání práce přímo v podnicích, využívání inzerce, podnikání kroků pro založení vlastní firmy, podání žádosti o pracovní povolení a licence nebo hledání zaměstnání jiným způsobem.



Predikce nezaměstnanosti v Plzeňském kraji

B. od 3. čtvrtletí 2004 (byla zavedena nová metodika výpočtu míry nezaměstnanosti srovnatelná s pravidly EU)

- **Čítatel:** v čitateli je uveden pouze počet tzv. **dosažitelných neumístěných uchazečů o zaměstnání**. Jsou to všichni uchazeči o zaměstnání podle zákona o zaměstnanosti č. 435/2004 §24, kteří mohou ihned nastoupit do zaměstnání. (Jsou započítáváni jen uchazeči, kteří mohou bezprostředně do 14 dnů nastoupit do zaměstnání.)
- **Jmenovatel:** „pracovní síla“ která je sestavena vzorcem jako součet **zaměstnaných z VŠPS, počet pracujících cizinců, počet cizinců s živnostenským oprávněním a registrovaní uchazeči o zaměstnání evidovaných na ú.p.**, (vše klouzavý roční průměr).

2.3 Popis současného stavu a východisek řešení

Po popisu zdrojů a dat charakterizujte současnou situaci a pokuste se prozkoumat možné vztahy mezi zkoumanými ukazateli. Za tímto účelem můžete využít statistických metod, se kterými jste se setkali v předmětu Pravděpodobnost a statistika KMA/PSA či matematické modely v ekonometrii KMA/MME.

Popis současného stavu:

- Jak už bylo zmíněno dříve, tento model bude kauzálním predikčním modelem, založeným na velikostech toků nezaměstnaných mezi jednotlivými okresy. Model bude sestavován nad daty z časového intervalu 1.1.1999 až 31.12.2006, získaných ze zdroje MPSV (www.mpsv.cz). Data jsou získávána měsíčně a stejně tak predikce by měla být provedena po měsících.
- Vývoj nezaměstnanosti ve všech okresech je stabilní, vykazující cyklické sezónní výkyvy, jak je patrné z grafu viz Graf 1.

Tento model by měl predikovat počet nezaměstnaných od ledna 2007 do prosince 2010. Práce je vypracována na začátku listopadu 2007, takže bude možné zčásti prověřit, zda model funguje.

Z grafu (Graf 1) je patrné, že průběh nezaměstnaností v jednotlivých okresech je značně podobný. Závislost také potvrzují hodnoty výběrových korelačních koeficientů³ nezaměstnaností jednotlivých okresů, které jsou ve většině případů vyšší než 0,5. Tyto koeficienty jsou uvedeny v „Tabulka 1“:

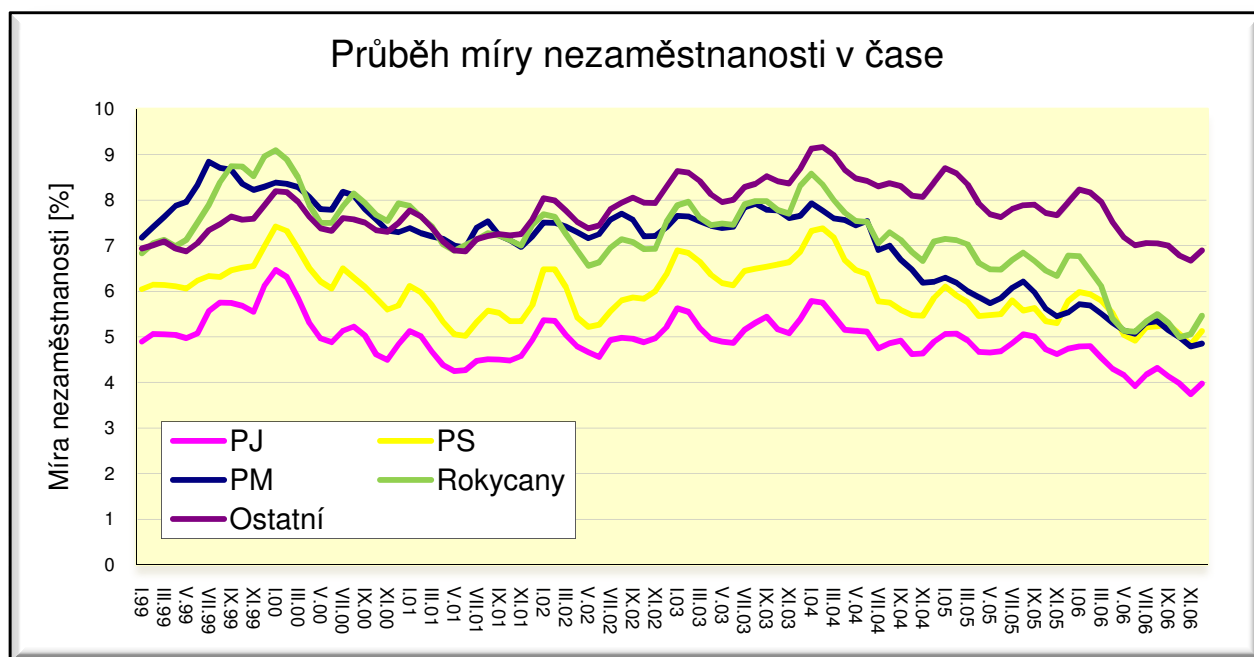
	PM	PJ	PS	Rokycany
PM	1	0,678344	0,703789	0,886262
PJ	0,678344	1	0,379671	0,586898
PS	0,703789	0,379671	1	0,827585
Rokycany	0,886262	0,586898	0,827585	1

Tabulka 1 – výběrová korelační matice nezaměstnanosti

³ Problematika korelace a výpočtu výběrového korelačního koeficientu je uvedena viz 1. přednáška – *mrf-1-regrese1.doc*



Predikce nezaměstnanosti v Plzeňském kraji



Graf 1

Tato korelovanost může být ale klamná, tzn. vyvolána také svazující podmínkou $\sum u_i = 100\%$, kde u_i je nezaměstnanost v okrese [%] a $i = 1 \dots$ počet okresů v ČR, či společnou závislostí nezaměstnanosti na inflaci – popsáno viz 2.1. Nicméně s ohledem na to, že dané okresy spolu navzájem sousedí, předpokládáme, že tady vzájemná závislost vývoje počtu nezaměstnaných v jednotlivých okresech existuje. Na základě této úvahy jsme pro úplnost zjistili také hodnoty výběrových korelačních koeficientů počtů nezaměstnaných jednotlivých okresů a otestovali jejich nenulovost – viz *Excel – list počet nezaměstnaných*.

Nyní víme, že má smysl sestavovat lineární predikční model počtu nezaměstnaných jednoho okresu na základě vývoje počtu nezaměstnaných v jiných okresech.

Aby byla zachována reálnost kauzálního modelu, bude zohledněno zjištění toků nejen mezi okresy, které jsou předmětem našeho zájmu, ale i jejich těsně sousedícím okolím – v práci nazýváno jako *Ostatní*. Data o tom, kolik zaměstnaných v okrese pochází z okresu jiného, či kolik zaměstnaných lidí z tohoto okresu je zaměstnáno v jiném okrese, však nejsou k dispozici.

Východiska:

- Můžeme sledovat počet volných míst v okrese v závislosti na počtu volných míst ve zbylých okresech a množství nezaměstnaných v okresech
- Můžeme přímo sledovat vývoj počtu nezaměstnaných v jednotlivých okresech v závislosti na počtu nezaměstnaných v ostatních okresech

Pro náš model zvolíme 2. variantu a to ze dvou důvodů:

Druhá varianta je jednodušší – je potřeba určit méně koeficientů a zároveň korelační koeficienty počet nezaměstnaných – počet nezaměstnaných vykazují vyšší hodnoty než vztah počet nezaměstnaných - volná místa, tzn. lépe popisuje vztahy mezi okresy – viz *Excel – list Počet volných míst*.



Predikce nezaměstnanosti v Plzeňském kraji

2.4 Popis možných variant řešení

Po identifikaci možných interpretovatelných vztahů mezi vhodnými ukazateli navrhněte možná řešení modelu. Zamyslete se nad více možnostmi řešení.

Jak bylo uvedeno v předchozí kapitole, budeme modelovat toky sledováním vývoje počtu nezaměstnaných v daném okrese v závislosti na počtech nezaměstnaných ve zbylých okresech. Tím je vyřešen problém modifikace toků. Hlavním cílem této studie je ale predikce velikosti nezaměstnanosti na následující 3 roky na základě sestrojení kauzálního modelu toků.

1. varianta řešení:

sestavit lineární regresní kauzální model ve tvaru:

$$(2.2) \quad \begin{bmatrix} N_{PM\ t} \\ N_{PJ\ t} \\ N_{PS\ t} \\ N_{Rokycany\ t} \\ N_{Ostatní\ t} \end{bmatrix} = A \begin{bmatrix} N_{PM\ t-1} \\ N_{PJ\ t-1} \\ N_{PS\ t-1} \\ N_{Rokycany\ t-1} \\ N_{Ostatní\ t-1} \end{bmatrix}, \text{ kde}$$

$N_{Okres\ t}$.. počet nezaměstnaných v Okrese v čase t

A .. matice regresních koeficientů

Hodnoty počtu nezaměstnaných $N_{Okres\ t}$ budou generovány na základě zpožděných hodnot všech okolních okresů. Regresní koeficienty obsažené v matici A charakterizují velikost toků mezi okresy. Modifikací tohoto modelu můžeme získat vztah generující velikost pracovní síly v okrese a tím přibližně určit velikost nezaměstnanosti. Vše je realizováno na základě lineární regrese.

2. varianta řešení

Sestavit model, ve kterém bude k predikci hodnot nezaměstnanosti využita časová řada. Časová řada bude sestavena pouze pro jeden okres. Pro zbylé okresy bude na základě dat, která jsou k dispozici odhadnuta soustava kauzálních rovnic generující hodnotu jejich nezaměstnanosti v závislosti na výpočtech časové řady u jednoho okresu. Model by pak vypadal následujícím způsobem:

1. Pro počet nezaměstnaných v okrese PM (N_{PM}), bude sestavena časová řada.

2. Kauzální model:

$$N_{PJ} = a_1 N_{PM} + b_1$$

$$N_{PS} = a_2 N_{PM} + b_2$$

$$N_{Rokycany} = a_3 N_{PM} + b_3$$

.....

Pozn.: V kauzálním modelu jsou počty nezaměstnaných N ze stejného času, není mezi nimi zpoždění, na rozdíl od předchozího modelu.

2.5 Vybraná varianta řešení

Z navržených možností řešení vyberte vhodnější a uveďte na základě jakých úvah jste tak učinili.

Pro náš model bude vhodnější 1. varianta řešení, která splňuje veškeré požadavky zadání – lépe specifikuje toky mezi okresy. Zároveň je tato varianta méně složitá z výpočetního hlediska. Při



Predikce nezaměstnanosti v Plzeňském kraji

generování časové řady by musela být uvažována i sezónnost a tím složitost druhé varianty značně převyšuje složitost varianty první.

3 Popis a podrobná dokumentace vybraného řešení

Nyní vybraný model identifikujte, kvantifikujte, verifikujte a aplikujte nad reálnými daty. Za tímto účelem využijte statistických a matematických nástrojů a znalostí z předmětů KMA/PSA, KMA/MME, KIV/MRF.

3.1 Popis dat

Data, nad kterými bude model vytvářen, jsou z intervalu: 31.1.1999 – 31.12.2006. K dispozici máme následující strukturu dat, v níž uvádíme jen položky užitečné pro tuto práci:

- Pracovní síla (klouzavý průměr za poslední 4 čtvrtletí)
- Neumístění ke konci minulého měsíce
- Nově hlášení neumístění
- Vyřazení ve sledovaném měsíci
- Volná pracovní místa ke konci sledovaného měsíce
- Přírůstek nezaměstnanosti
- Nezaměstnanost

Jak bylo zmíněno na začátku, budeme se zabývat toky nezaměstnaných mezi okresy PM, PJ, PS, Rokycany. Aby byl tento model zcela objektivní, musely být uvažovány i další nejbližší okolní regiony, reprezentované pod názvem *Ostatní*. Patří sem (podle NUTS 3):

- Praha
- Středočeský kraj
- Jihočeský kraj
- Plzeňský kraj
- Karlovarský kraj
- Ústecký kraj

Problém se změnou uspořádání

Od roku 1.1.2000 bylo v ČR zavedeno nové regionální uspořádání NUTS 3. Proto musela být data pro rok 1999 modifikována tak, aby byla kompatibilní. Transformace:

Jihočeský kraj - Pelhřimov = Jihočeský kraj (nový)

Severočeský kraj - Česká Lípa - Jablonec nad Nisou - Liberec = Ústecký kraj

Západočeský kraj = Plzeňský kraj + Karlovarský kraj



Predikce nezaměstnanosti v Plzeňském kraji

Problém se změnou metodiky měření v průběhu sledovaného intervalu

Jak již bylo zmíněno v kapitole 2.2, metodika měření nezaměstnanosti se změnila přechodem z června 2004 na červenec 2004. Mezi nezaměstnané se začali započítávat jen osoby, které jsou dosažitelné – tj. jsou schopny nastoupit do zaměstnání během 14dnů. Tím došlo k poklesu počtu evidovaných nezaměstnaných ve sledovaných oblastech v průměru o 8,3 procentního bodu. Do pracovní síly jsou nově započítáváni také pracující cizinci a cizinci s živnostenským oprávněním to zvýšilo velikost pracovní síly ve sledovaných oblastech v průměru o 2,55 procentního bodu. Celkově modifikace metodiky měření nezaměstnanosti vykazuje míry nezaměstnanosti v průměru o 0,7 procentního bodu nižší oproti hodnotě míry nezaměstnanosti zjišťované starší metodikou.

Pro účely našeho kauzálního modelu budeme pracovat s počty nezaměstnaných jedinců. Abychom data sjednotili, snížíme hodnoty počtu uchazečů z časového intervalu od ledna 1999 do června 2004 oproti původním o průměrné procento změny z původní hodnoty ve sledované oblasti - viz Excel – list Modifikovaný počet uchazečů.

3.2 Specifikace ekonometrického modelu

A. Lineární regresní kauzální model počtu nezaměstnaných

Lineární regresní kauzální model má tvar:

$$(3.1) \quad \begin{bmatrix} N_{PM_t} \\ N_{PJ_t} \\ N_{PS_t} \\ N_{Rokycany_t} \\ N_{Ostatní_t} \end{bmatrix} = A \begin{bmatrix} N_{PM_{t-1}} \\ N_{PJ_{t-1}} \\ N_{PS_{t-1}} \\ N_{Rokycany_{t-1}} \\ N_{Ostatní_{t-1}} \end{bmatrix} + u,$$

- kde
- $A \in \mathbb{R}^{5 \times 5}$ je matice regresních koeficientů modelu, které určujeme,
 - $u \in \mathbb{R}^5$ je náhodná složka,
 - $t \in \mathbb{R}$ je čas udávaný v měsících,
 - N_{Okres_t} je počet nezaměstnaných v Okresu v čase t .

Matici regresních koeficientů získáme složením koeficientů jednotlivých regresních rovnic. Pro ilustraci je uveden následující příklad 1. regresní rovnice:

$$N_{PM_t} = a_{11}N_{PM_{t-1}} + a_{12}N_{PJ_{t-1}} + a_{13}N_{PS_{t-1}} + a_{14}N_{Rokycany_{t-1}} + a_{15}N_{Ostatní_{t-1}} + u_1$$

Jak je patrné z modelu, jedná se o model lineární regrese bez absolutního členu. Absolutní člen nezapočítáváme vzhledem k problémům⁴, které by nastaly při převodu modelu na model s nezaměstnaností, jak uvidíme dále.

⁴ Problematika lineární regrese, vlastnosti regresní matice a vliv absolutního členu na výsledná rezidua a koeficient determinace je uvedena v přednášce *mrf-1-regrese1.doc*.



Predikce nezaměstnanosti v Plzeňském kraji

B. Model pro predikci nezaměstnanosti založený na lineárním regresním kauzálním model počtu nezaměstnaných

Výše uvedený regresní model umožní predikovat počet nezaměstnaných příštích měsíců. K přibližné predikci nezaměstnanosti je zapotřebí znát velikost pracovní síly v okrese – značíme E_{Okresu} . K tomuto problému můžeme přistoupit dvěma způsoby:

1. hodnoty E_{Okresu} budou v čase proměnné a určíme je lineární, či jinou regresí
2. hodnoty E_{Okresu} budou fixní a určíme je modifikací již dříve použitého modelu.

Protože přesná predikce ukazatele nezaměstnanosti není hlavním cílem této práce, pro orientační výpočet bude postačující druhý způsob. Zároveň prokáže mnohostranné využití tohoto modelu. Model 3.1. (v něm) modifikujeme následujícím způsobem:

$$N_{PM_t} = a_{11} \frac{N_{PM_{t-1}}}{E_{PM}} E_{PM} + a_{12} \frac{N_{PJ_{t-1}}}{E_{PJ}} E_{PJ} + a_{13} \frac{N_{PS_{t-1}}}{E_{PS}} E_{PS} + a_{14} \frac{N_{Rokycany_{t-1}}}{E_{Rokycany}} E_{Rokycany} + a_{15} \frac{N_{Ostatní_{t-1}}}{E_{Ostatní}} E_{Ostatní}$$

použijeme substituci $\alpha_{11} = a_{11} E_{PM}$, $\alpha_{12} = a_{12} E_{PJ}$, atd.

Získáme tak novou regresní rovnici počtu nezaměstnaných s vysvětlujícími proměnnými reprezentovanými nezaměstnaností jednotlivých okresů z minulého období. Tvar rovnice je:

$$N_{PM_t} = \alpha_{11} \frac{N_{PM_{t-1}}}{E_{PM}} + \alpha_{12} \frac{N_{PJ_{t-1}}}{E_{PJ}} + \alpha_{13} \frac{N_{PS_{t-1}}}{E_{PS}} + \alpha_{14} \frac{N_{Rokycany_{t-1}}}{E_{Rokycany}} + \alpha_{15} \frac{N_{Ostatní_{t-1}}}{E_{Ostatní}}$$

$$N_{PM_t} = \alpha_{11} v_{PM_{t-1}} + \alpha_{12} v_{PJ_{t-1}} + \alpha_{13} v_{PS_{t-1}} + \alpha_{14} v_{Rokycany_{t-1}} + \alpha_{15} v_{Ostatní_{t-1}}$$

kde $v_{PM_{t-1}}$ je nezaměstnanost v okrese Plzeň Město v čase $t-1$.

Pro tento regresní model určíme znovu hodnoty koeficientů α_{ij} . Jejich hodnoty pak následujícím způsobem porovnáme s hodnotami modelu předchozího a získáme tak hodnoty E_{Okres} .

$$\alpha_{11} = a_{11} E_{PM} \rightarrow E_{PM} = \frac{\alpha_{11}}{a_{11}}$$

$$(3.2) \quad \alpha_{ii} = a_{ii} E_{Okres}, \quad \alpha_{ii} = a_{ii} E_{Okres} \rightarrow E_{OkresM} = \frac{\alpha_{ii}}{a_{ii}}$$

Pro odhad E_{Okres} budeme uvažovat pouze koeficienty a_{ii} , které jsou statisticky významné⁵, tzn. mají vliv na počet nezaměstnaných. Hodnoty E určujeme na základě diagonálních prvků, protože ty nesou o velikosti pracovní síly v okrese největší informaci.

C. Modelování velikosti toků nezaměstnaných mezi okresy

Model (3.1) dále umožňuje určit odhad velikosti toků mezi jednotlivými okresy:
Platí:

$$(3.3) \quad N_{O_i_t} = N_{O_{i-1}} + \Delta_{O_i,O_1} + \Delta_{O_i,O_2} + \Delta_{O_i,O_3} + \Delta_{O_i,O_4}$$

⁵ Více o statistické významnosti odhadu jednotlivých regresních koeficientů je uvedeno viz přednáška *Mrf-1-regrese2.doc*



Predikce nezaměstnanosti v Plzeňském kraji

kde $\Delta_{oi,oj}$ reprezentuje tok z Okresu j do Okresu i. Pro $i=1$ platí:

$$N_{o1_t} = N_{o1_{t-1}} + \frac{\Delta_{oi,o1}}{N_{o1_{t-1}}} N_{o1_{t-1}} + \frac{\Delta_{oi,o2}}{N_{o2_{t-1}}} N_{o2_{t-1}} + \frac{\Delta_{oi,o3}}{N_{o3_{t-1}}} N_{o3_{t-1}} + \frac{\Delta_{oi,o4}}{N_{o4_{t-1}}} N_{o4_{t-1}}$$

$$N_{o1_t} = \left(1 + \frac{\Delta_{oi,o1}}{N_{o1_{t-1}}}\right) N_{o1_{t-1}} + \frac{\Delta_{oi,o2}}{N_{o2_{t-1}}} N_{o2_{t-1}} + \frac{\Delta_{oi,o3}}{N_{o3_{t-1}}} N_{o3_{t-1}} + \frac{\Delta_{oi,o4}}{N_{o4_{t-1}}} N_{o4_{t-1}}.$$

Tyto rovnice porovnáme s regresními koeficienty a_{ij} základního modelu a to následujícím způsobem:

$$(3.4) \quad a_{ii} = \left(1 + \frac{\Delta_{oi,o1}}{N_{o1_{t-1}}}\right), \quad a_{ij} = \frac{\Delta_{oi,oj}}{N_{oj_{t-1}}}, \text{ kde } j \neq i, \quad i = 1..5$$

Z těchto vztahů odhadneme toky $\Delta_{oi,oj}$ mezi okresy.

3.3 Kvantifikace ekonometrického modelu

A. Sestavení lineárního regresního modelu z kapitoly 3.2

Ke zjištění regresních koeficientů použijeme metodu nejmenších čtverců MNČ.

Předpoklady:

- 1) $E(u) = 0$
- 2) $E(uu') = \sigma^2 I$... požadavek na homoskedasticitu
- 3) N je nestochastická matice ... $E(N^T u) = 0$
- 4) N má plnou hodnotu = 5

N .. matice s daty nad kterými je model vytvářen (v našem modelu sloupce odpovídají jednotlivým zkoumaným oblastem, tj. 5 sloupců, celkem máme 96 x 5 hodnot)

Matici sestavíme složením získaných regresních rovnic pro jednotlivé okresy. Získané výsledky jsou uvedeny viz Tabulka 2, ve které jsou u jednotlivých modelů okresů uvedeny hodnoty koeficientů – a_1, \dots, a_5 . U jednotlivých regresních koeficientů je uveden 95% interval spolehlivosti $\langle a_i - s_{a_i} t_{\alpha/2}; a_i + s_{a_i} t_{\alpha/2} \rangle$. Kde s_{a_i} je standardní chyba pro jednotlivé regresní koeficienty a_i a $t_{\alpha/2}$ je kvantil Studentova rozdělení. (K tomu, abychom mohli použít tento test významnosti jednotlivých koeficientů, musíme pro náhodnou složku doplnit předpoklad, $N(0, \sigma^2)$ rozdělení.) Významné hodnoty koeficientů jsou vyznačeny hnědě. Dále jsou uvedeny hodnoty koeficientu determinace R^2 a rozptyl náhodné složky σ^2 (chyba střední regresní hodnoty).



Predikce nezaměstnanosti v Plzeňském kraji

Okres	a_1	a_2	a_3	a_4	a_5	R^2	σ^2
PM	1.055	-0,162	0.075	-0,222	0.001	0.999	179,65
Interval spolehlivosti	<0,956 1,154>	<-0.84 0,516>	<-0.732 0.36>	<-0,711 0.26>	<-0.003 0.005>		
PJ	0.007	0.938	-0.106	0.068	0.001	0.999	72.483
Interval spolehlivosti	<-0.033 0.047>	<0.66 1.21>	<-0.281 0.068>	<-0.129 0.266>	<-0.001 0.003>		
PS	0.010	0.273	0.681	0.122	0.000	0.999	96,25
Interval spolehlivosti	<-0.043 0.063>	<-0.090 0.637>	<0.449 0.912>	<-0.141 0.384>	<-0.002 0.002>		
Rokycany	0.044	0.043	-0.088	0.882	-0.0003	0.999	63.665
Interval spolehlivosti	<0.008 0.079>	<-0.197 0.284>	<-0.242 0.065>	<0.708 1,055>	<-0.001 0.002>		
Ostatní	1.094	5.33	-11.3	5.652	0.0595	0.999	4721.02
Interval spolehlivosti	<-1.512 3,7>	<-12,5 23,16>	<-22.67 0.073>	<-7.211 18,515>	<0.888 1,124>		

Tabulka 2

Z koeficientu determinace je patrné, že odhadnuté koeficienty lineárních regresních rovnic dobře popisují kauzální závislost počtu nezaměstnaných v jednom okrese na počtu nezaměstnaných v minulém období ve všech okresech. Nicméně je zde velké množství statisticky nevýznamných koeficientů. Proto byl sestaven nový lineární regresní model tak, aby obsahoval jen ty koeficienty, na nichž vývoj počtu nezaměstnaných skutečně záleží. (Vytvoříme regresní model jen nad sloupci dat, k nimž se vztahuje statisticky významná hodnota viz Tabulka 2). Získáme pak regresní model popsany následující maticí A. (V tomto modelu byla stejně jako v předchozím případě testována také významnost koeficientů.).

$$(3.5) \quad \text{Matice } A = \begin{bmatrix} 0,998 & 0 & 0 & 0 & 0 \\ 0 & 0,9997 & 0 & 0 & 0 \\ 0 & 0 & 0,9995 & 0 & 0 \\ 0,040 & 0 & 0 & 0,8502 & 0 \\ 0 & 0 & 0 & 0 & 1,001 \end{bmatrix}$$

Odhad chyby aproximace pomocí směrodatné odchylky chybové složky $S_e = \sqrt{\frac{SSE}{n-2}}$

PM = 177,97 PJ = 73,14 PS = 98,45 Rokycany = 62,92 Ostatní = 4 812

B. Sestavení regresního modelu s nezaměstnaností

$$\begin{bmatrix} N_{PM_t} \\ N_{PJ_t} \\ N_{PS_t} \\ N_{Rokycany_t} \\ N_{Ostatní_t} \end{bmatrix} = \hat{A} \begin{bmatrix} v_{PM_{t-1}} \\ v_{PJ_{t-1}} \\ v_{PS_{t-1}} \\ v_{Rokycany_{t-1}} \\ v_{Ostatní_{t-1}} \end{bmatrix} + u,$$

u .. náhodná složka, u které je předpokládáno $N(0, \sigma^2)$



Predikce nezaměstnanosti v Plzeňském kraji

Použijeme znovu metodu nejmenších čtverců a získáme:

Okres	α_1	α_2	α_3	α_4	α_5	R^2	σ^2
PM	463,44	681,28	-393,493	-189	392,59	0.998	305,72
Interval spolehlivosti	<315,95 611>	<318,95 1 043>	<-672,85 -114,13>	<-404,5 26,4>	<279,37 505,8>		
PJ	-110,48	428,26	-168,322	-60,01	222,21	0.998	104,31
Interval spolehlivosti	<-160,8 -60,1>	<304,63 551,9>	<-263,64 -73,0>	<-113,5 13,5>	<-263,64 -73>		
PS	-158,1	419,3	50,03	-64,5	184,18	0.9979	147,15
Interval spolehlivosti	<-229,13 -87,03>	<244,9 593,7>	<-84,43 184,48>	<-168,2 39,2>	<129,7 238,7>		
Rokycany	-64,03	148	-134,12	152,7	133,05	0.9982	98,88
Interval spolehlivosti	<-111,8 -16,28>	<30,8 265,2>	<-224,48 -43,76>	<83 222,4>	<96,43 169,67>		
Ostatní	-5283,65	10410,9	-9496,9	-2747,7	29729,2	0,99936	6174,03
Interval spolehlivosti	<-8264,9 -2302>	<3093,5 17728>	<-15138 -3855>	<-7099 1603>	<27442 32016>		

Tabulka 3

Opět model dobře popisuje data, jak je patrné z koeficientu determinace. Na základě intervalů spolehlivosti jednotlivých koeficientů byly vybrány sloupce se statisticky významnými koeficienty. Pro ně byl pak sestaven model nový. Výsledná matice koeficientů je

$$\hat{A} = \begin{bmatrix} 372,2 & 555 & -393,34 & 0 & 380,06 \\ -139,4 & 388,21 & -168,3 & 0 & 218,2 \\ -184,9 & 416,4 & 0 & 0 & 188,91 \\ -64 & 148 & -134,1 & 152,7 & 133,05 \\ -6609,7 & 8576,98 & -9494,98 & 0 & 29547,1 \end{bmatrix}$$

Nyní určíme velikost pracovní síly v jednotlivých okresech podle vztahu 4.2. $E_{OkresM} = \frac{\alpha_{ii}}{a_{ii}} * 100$

(násobeno 100 kvůli míře nezaměstnanosti, která je udávána v %). Hodnoty jsou uvedeny viz Tabulka 4. Pro ověření reálnosti vypočítaných hodnot je v tabulce uvedena průměrná hodnota a medián reálných dat, nad kterými byl model sestavován. Současně je zkoumána relativní chyba.

Okres	PM	PJ	PS	Rokycany	Ostatní
E_{Okresu}	37278,5	38821,7	0	17958,6	2951684,8
Průměrná hodnota E_{Okresu}	88453,8	33826,2	38586,5	23207,4	2254243,9
Diference = prům.h - E_{Okresu}	51175,3	-4995,5	38586,5	5248,9	-697440,9
Relativní chyba = diference/prům.h	57,86%	-14,77%	100,00%	22,62%	-30,94%
Medián E_{Okresu}	88123	33684,5	38369	23096	2229747
Diference = medián - E_{Okresu}	-330,8	-141,7	-217,5	-111,4	-24496,9
Relativní chyba diference/medián	-0,38%	-0,42%	-0,57%	-0,48%	-1,10%
Hodnota ke 31.11.2003	92 746	37 067	40 192	24 518	2 394 605
Diference 31.11.03 - E_{Okresu}	55 468	-1 755	40 192	6 559	-557 080
Relativní chyba diference/h. ke 31.1.03	0,6	-0,05	1	0,3	-0,2

Tabulka 4



Predikce nezaměstnanosti v Plzeňském kraji

Z tabulky je patrné, že daný model pro zjištění střední velikosti pracovní síly v okrese se dá použít pro okres Plzeň jih. V případě ostatních okresů je relativní chyba reálných dat od modelovaných hodnot vyšší než 20% průměrné hodnoty. Nejvyšší diferenci vykazuje PS, u kterého je střední hodnota velikosti pracovní síly rovna nule, neboť koeficient $\alpha_{3,3} = 0$.

C. Sestavení modelu pro výpočet odhadu toků

Pro odhad velikosti toku nezaměstnaných v čase t vyjdeme ze vztahu (3.4)

$$a_{ii} = \left(1 + \frac{\Delta_{Oi,Oi}}{N_{Oi,t-1}} \right), \quad a_{ij} = \frac{\Delta_{Oi,Oj}}{N_{Oj,t-1}},$$
 z něho pak pomocí získaných regresních koeficientů dojdeme k rovnicím popisujícím tok nezaměstnaných z okresu j do okresu i v čase t :

$$N_{Oi,t-1}(a_{ii} - 1) = \Delta_{ii}, \quad N_{Oj,t-1}a_{ij} = \Delta_{ij}$$

Rovnice toků nezaměstnaných mezi okresy jsou:

PM

$$N_{PM,t-1}(-0,0016) = \Delta_{PM,PM},$$

PJ

$$N_{PJ,t-1}(-0,00003) = \Delta_{PJ,PJ},$$

PS

$$N_{PS,t-1}(-0,0005) = \Delta_{PS,PS},$$

Rokycany

$$N_{Rokycany,t-1}(-0,1498) = \Delta_{Rokycany,Rokycany}, \quad N_{PM,t-1}(0,04033) = \Delta_{Rokycany,PM}$$

Ostatní

$$N_{Ostatní,t-1}(-0,00103) = \Delta_{Ostatní,Ostatní},$$

3.4 Verifikace modelu

▪ Ekonomická verifikace

Matice A má statisticky významné prvky vždy na diagonále, což potvrzuje vlastnost, že vývoj nezaměstnanosti v okrese nejvíce ovlivňuje předchozí počet nezaměstnaných v tomto okrese. Prvky mimo diagonálu ukazují, že se mezi okresy nachází tok. Pokud je číslo záporné, znamená to, že osoby dostali v j -tém okrese zaměstnání a v i -tém (domácím) okrese byl snížen počet nezaměstnaných. Opačně, kladné znaménko znamená, že osoby v j -tém okrese ztratily zaměstnání a tak narostla nezaměstnanost v domácím i -tém okrese. Dosažené hodnoty tedy z ekonomického hlediska dávají smysl.

▪ Statistická verifikace

Statistická významnost a tedy i skutečná závislost vysvětlující proměnné na vysvětlované byla prověřena už při sestavování regresního modelu. Model tedy obsahuje jen statisticky významné regresní koeficienty.

Pro takto upravený model byly vypočítány koeficienty determinace R^2 , koeficient vícenásobné determinace R_v^2 (zohledňující počet odhadovaných regresních koeficientů), F statistika udávající významnost R^2 . Jejich hodnoty jsou uvedeny viz Tabulka 5.

$$R^2 = 1 - e'e/N'N, \text{ kde } e = N_{\text{odhad}} - N_{\text{skutečné}},$$



Predikce nezaměstnanosti v Plzeňském kraji

$$R^2_v = R^2 - \frac{k-1}{n-k} (1 - R^2)$$

$$F = \frac{R^2}{1 - R^2} \frac{n - k}{k - 1} \sim F^*(4, 91) = 2,472 \text{ při } 95\% \text{ hladině významnosti}$$

k .. počet regresních koeficientů

n .. rozsah výběru dat

Okres	PM	PJ	PS	Rokycany	Ostatní
R^2	0,9992	0,9992	0,9993	1,0011	0,9995
R^2_v	0,9992	0,9992	0,9993	1,0011	0,9995
F	28423,9	29172,3	32531,5	-20721,7	46690,3

Tabulka 5

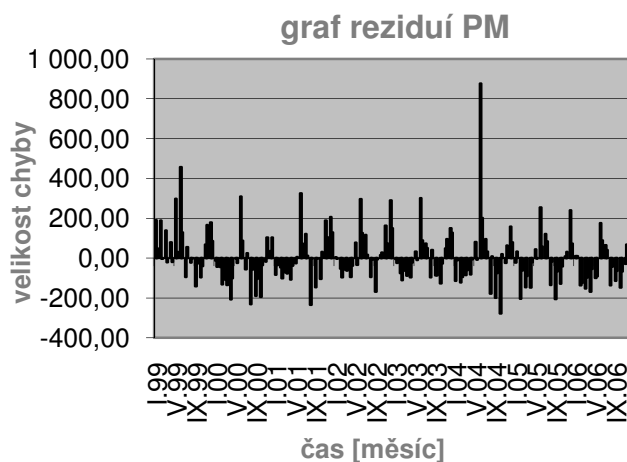
Ze získaných hodnot koeficientů determinace je patrné, že model pro výpočet počtu nezaměstnaných se velmi dobře shoduje s daty – vysoké hodnoty koeficientu F prokazují fakt, že koeficienty determinace jsou statisticky významné a že mezi vysvětlujícími proměnnými a vysvětlovanou proměnnou existuje významný vztah.

▪ *Ekonometrická verifikace*

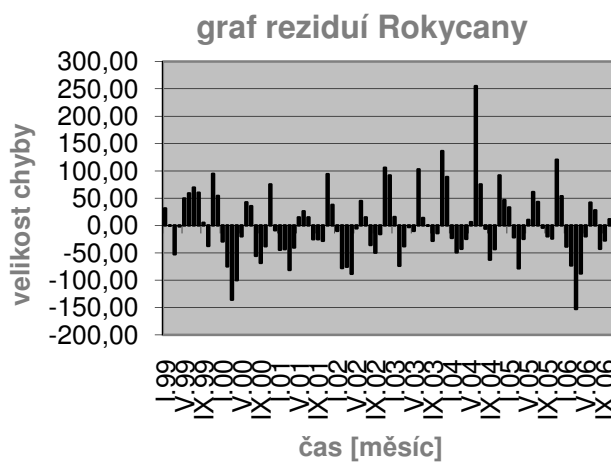
V této části budeme prověřovat předpoklady použití metody nejmenších čtverců uvedených v části 3.3

1/ prověření multikolinearity – sloupce matice dat, nad kterými byl model sestavován jsou lineárně nezávislé. To také potvrzuje případ, kdy několik hodnot dat bylo změněno. Koeficient determinace byl stále stejně vysoký, regresní koeficienty se změnily jen málo (úměrně změně dat). To zároveň potvrzuje robustnost sestaveného modelu.

2/ test heteroskedasticity - využijeme grafického znázornění reziduí. Viz Graf 2 až Graf 6 je patrné, že rezidua nevykazují žádnou funkční závislost nárůstu či poklesu svých velikostí, tzn. byl potvrzen předpoklad homoskedasticity.



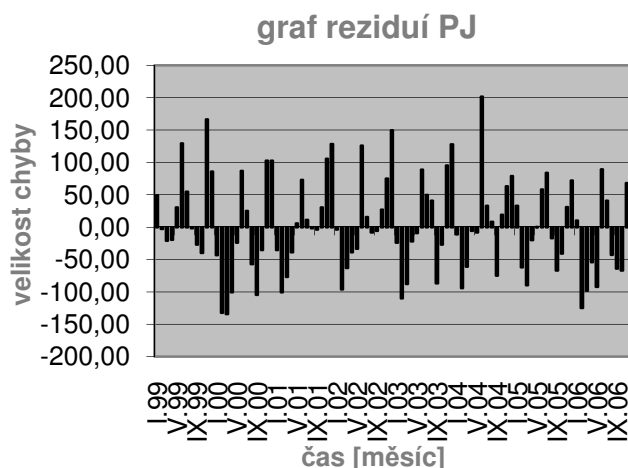
Graf 2



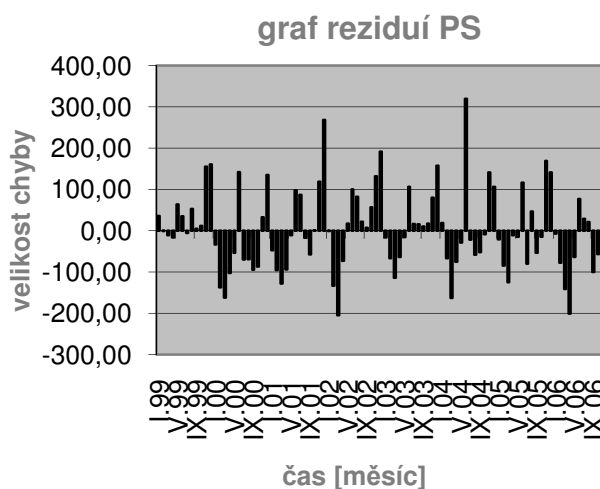
Graf 3



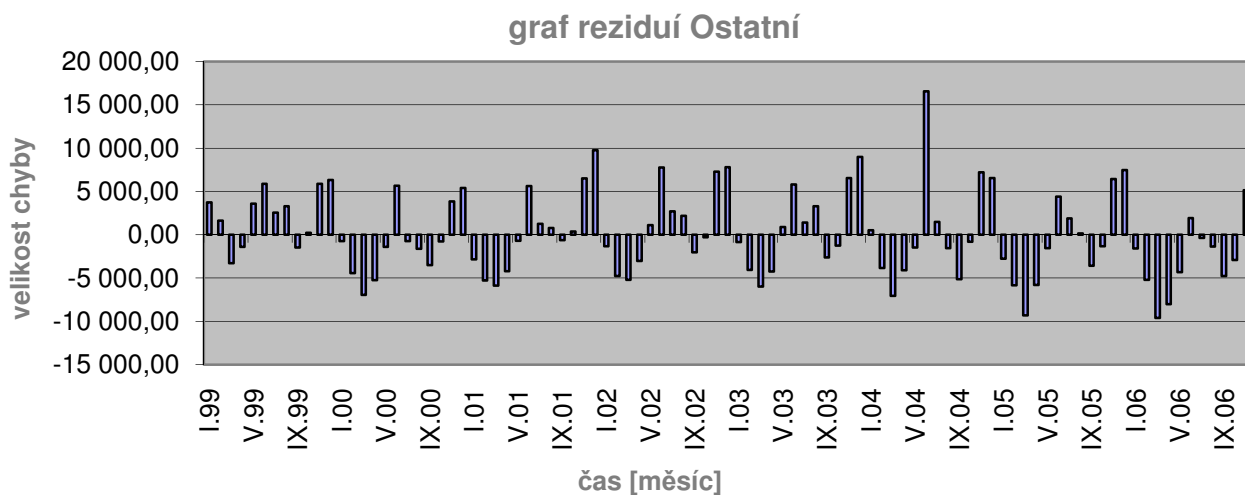
Predikce nezaměstnanosti v Plzeňském kraji



Graf 4



Graf 5



Graf 6

3/ test autokorelace náhodných složek – byl proveden D-W test na existenci autokorelace 1. řádu, kterým byly získány následující hodnoty:

$dw = [1,304 \quad 1,228 \quad 1,246 \quad 1,168 \quad 1,123]$. Všechny hodnoty jsou z intervalu (0;4) -> korelace prvního řádu náhodných složek se nevyskytuje.

3.5 Praktické využití – prognózování

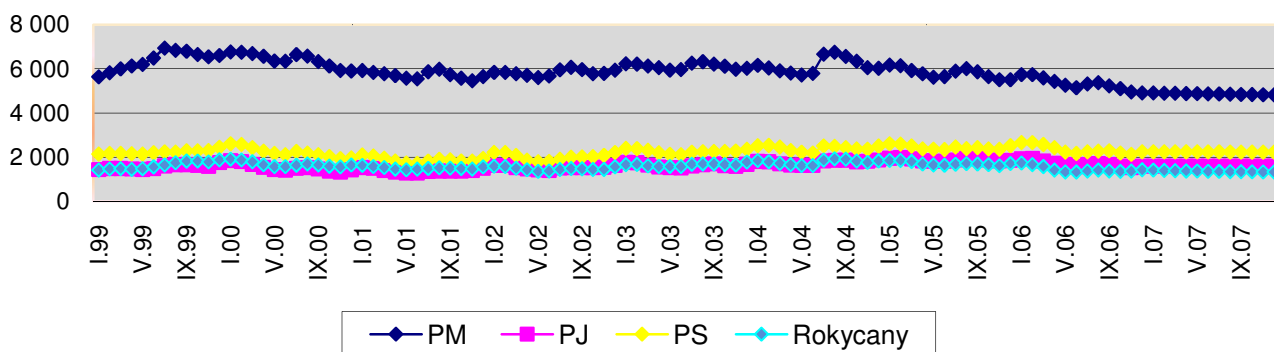
Po samotném sestavení a ověření správnosti výsledného modelu statistickými metodami využijte výsledný model k účelu, ke kterému byl stvořen, a aplikujte ho - v tomto případě provedeme prognózu.

Po použití modelu (3.1) a matice A (3.5) získáme hodnoty počtu nezaměstnaných podle jednotlivých okresů uvedené viz list Prognóza v příloženém Excelovském souboru. Graficky je prognóza spolu s původními reálnými daty, nad kterými byl model sestavován, zobrazena v níže uvedených grafech Graf 7 až Graf 11.

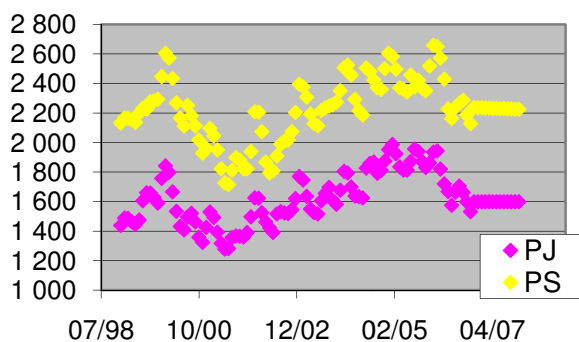


Predikce nezaměstnanosti v Plzeňském kraji

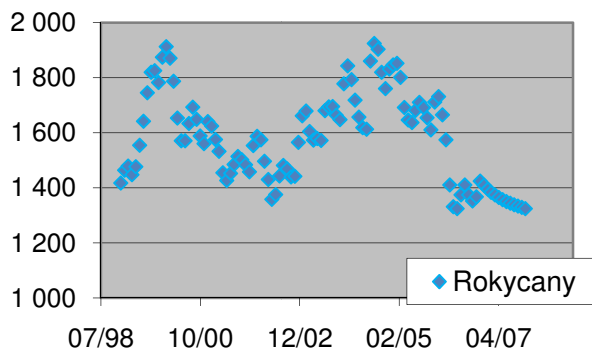
Hodnoty počtů nezaměstnaných s předpovědí bez ostatních okresů



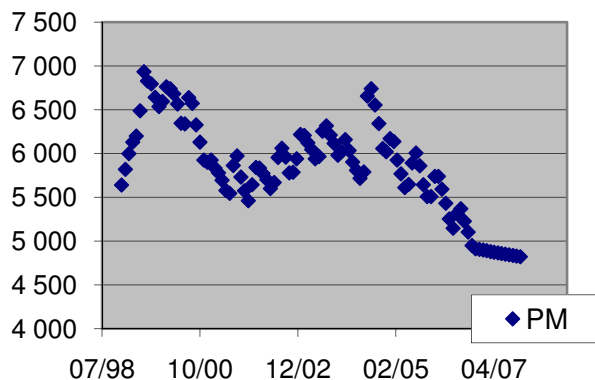
Graf 7



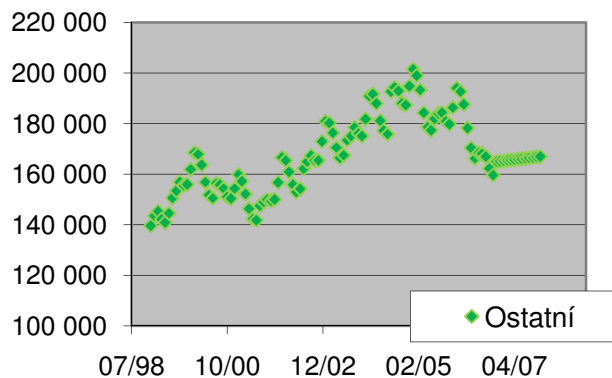
Graf 8



Graf 9



Graf 10



Graf 11

Průběh toků je vypočítán na základě modelu 3.3 C. Výsledky jsou uvedeny v Excelu - list Prognóza..

Z hodnot prognózy počtu nezaměstnaných je patrné, že počet nezaměstnaných v PM bude po celé 3 roky mírně klesat. Tento pokles, by dle našeho modelu neměl být způsoben vnějším tokem, nýbrž odtokem nezaměstnaných v rámci okresu. Celkový tok je predikován na 8 osob měsíčně (toto číslo je definováno rozdílem přítok – odtok v rámci okresu). Průběh počtu nezaměstnaných v okresech PJ i PS je také mírně klesající až konstantní. Vypočítaný tok je téměř nulový a je realizován pouze



Predikce nezaměstnanosti v Plzeňském kraji

v rámci domácího okresu. V Rokycanech dojde k úbytku počtu nezaměstnaných a to hlavně díky zápornému toku (nezaměstnaní např. najdou práci) v rámci domácího okresu. U tohoto okresu byl zjištěn významný kladný tok Rokycany – PM, který snižuje velikost úbytku počtu nezaměstnaných. Regiony Ostatní nejsou předmětem odhadu této práce, nicméně jejich průběh byl z důvodu realizace predikce nutný. (Počet nezaměstnaných se v okolních regionech zvyšuje, tok existuje pouze vnitřní.)

Pohled na Plzeňský region celkově:

Průběh počtu nezaměstnaných v Plzni a okolí (uvažovány všechny 4 zkoumané okresy) bude stabilní s klesajícím trendem. K největšímu úbytku počtu nezaměstnaných dojde v období 1/07 – 12/09 v okresech Plzeň město a Rokycany. Zbylé okresy v tomto období zaznamenají stagnaci či velice mírný pokles.

4 Diskuse

Na závěr konfrontujte výsledky modelu s realitou.

4.1 Diskuse nesouladů mezi technickým a fundamentálním modelem

Kauzální model toků, který jsme získali, nebude dávat výsledky, které by měsíčně přesně predikovaly počet nezaměstnaných díky tomu, že nezachycuje sezónnost, která se v počtu nezaměstnaných projevuje. (Jak je patrné viz Graf 1, největší je nezaměstnanost v lednu – únoru, srpnu – září, nejmenší v květnu a říjnu) Nicméně by měl poskytovat dobrou roční (průměrnou) predikci vývoje počtu nezaměstnaných i tendence vývoje. Model odhalil jen málo toků mezi okresy. To může být způsobeno nedostatečným množstvím bazických dat, nad kterými byl model sestaven.

4.2 Kritika dat

V této části se zaměříme na kritiku predikovaných dat. Porovnáme odhadovaná a skutečná data pro počet nezaměstnaných v jednotlivých okresech.

	PM	Odhad	PJ	Odhad	PS	Odhad	Rokycany	Odhad
01/07	5 215	4 904	1 439	1 600	2 391	2 238	1 467	1 409
02/07	5 076	4 897	1 374	1 600	2 337	2 237	1 432	1 396
03/07	4 907	4 889	1 261	1 600	2 137	2 236	1 332	1 384
04/07	4 669	4 881	1 200	1 600	1 996	2 235	1 296	1 374
05/07	4 608	4 874	1 155	1 600	1 947	2 234	1 247	1 365
06/07	4 699	4 866	1 125	1 600	1 932	2 233	1 217	1 357
Průměr	4 767	4 870	1 232	1 600	2 027	2 233	1 257	1 365
Chyba [% odhadu]		2%		23%		9%		8%

Tabulka 6

Jak je vidět, reálná data se od odhadovaných hodnot liší jen málo, vezmeme-li v úvahu vlivy sezónnosti, jimž průběh reálných dat odpovídá a které náš model, jak již bylo zmíněno, nepopisuje. Pro lepší zjištění shody modelu s daty byly vyjádřeny aritmetickými průměry průměrné hodnoty počtu nezaměstnaných za dané období pro daný okres a to jak pro hodnoty reálné tak i modelové. Následně pak byla vyjádřena chyba modelu od reality. Bylo tak zjištěno, že modely pro PM, Rokycany a PS popisují průběh nezaměstnanosti s chybou nepřesahující 10 procent odhadované hodnoty. Nejhůře skončil model pro PJ, v němž je odhadovaný počet nezaměstnaných vyšší než skutečnost. Vzhledem k malému množství reálných ověřovacích dat se ale může jednat jen o krátkodobý výkyv.



Predikce nezaměstnanosti v Plzeňském kraji

5 Zdroje dat

Nezapomeňte v textu důsledně uvádět odkazy na zdroje dat!!!

- [1] Metodika ČSÚ: [http://www.czso.cz/csu/2007edicniplan.nsf/t/55002D4D5B/\\$File/310107q2-2.pdf](http://www.czso.cz/csu/2007edicniplan.nsf/t/55002D4D5B/$File/310107q2-2.pdf)
- [2] Zdroj dat o počtu nezaměstnaných: <http://portal.mpsv.cz/sz/stat/nz/mes>